A Delegation Approach to Regulating Hiring Discrimination*

Paolo Martellini
University of Wisconsin Madison

Guido Menzio
NYU and NBER

December 2023

Abstract

We approach the design of anti-discriminatory labor market regulation as a delegation problem. A private firm (the agent) is repeatedly faced with the opportunity of hiring one among several applicants to fill its vacancies. The firm is biased against applicants from some demographic group, and it is neutral towards applicants from some other group. Applicants differ not only with respect to their demographic characteristics, but also with respect to the idiosyncratic quality of their match with firm. A benevolent and unbiased labor market authority (the principal) enacts a hiring regulation (a direct-revelation mechanism without transfers) in order to reduce the impact of the firm's bias on its hiring behavior. The hiring regulation is constrained by the fact that the quality of the match between any particular applicant and the firm is privately observed by the firm. We characterize the optimal mechanism.

JEL Codes: D82, J71.

Keywords: Mechanism Design, Delegation, Labor Discrimination.

---

*Martellini: Department of Economics, University of Wisconsin Madison, 1180 Observatory Drive, Madison, WI 53706 (email: martellini@wisc.edu). Menzio: Department of Economics, New York University, 19 West 4th Street, New York, NY 10012 (email: gm1310@nyu.edu). We previously circulated the manuscript as "The Dean and the Chair." We are grateful to Manuel Amador, Mark Aguiar, Jarda Borovicka, Borys Grochulski, Marina Halac, Oleg Itskhoki, Andreas Kleiner, Natalia Kovrijnykh, Jennifer La'O, Elliot Lipnowski, Lee Ohanian, Russell Wong, and Pierre Yared for helpful suggestions. We are also thankful to our audiences at the NBER Summer Institute, the SITE Workshop on Dynamic Contracts, Columbia, ASU, EIEF, CMU, FRB Chicago and FRB Richmond.

1 Introduction

Labor market legislation typically forbids firms from explicitly discriminating applicants from different demographic groups by offering wages that depend on an applicant's demographic characteristics. Firms, indeed, are not allowed to post a job opening that offers 50,000$ a year to a man, and 40,000$ a year to a woman. Such legislation, however, is unlikely to be effective at curbing discrimination, as long as firms are free to choose which applicants to hire and which applicants to ignore. For instance, a firm that posts a job opening offering 50,000$ a year to any applicant could still act on its gender bias by setting different standards for hiring men or women. A successful anti-discrimination legislation needs to include some restriction on the firms' hiring behavior. Obviously, the extent of hiring restrictions is limited by the fact that a firm has private information about how suitable different applicants are to fill its job openings.

In this paper, we approach the design of anti-discriminatory hiring regulation as an optimal delegation problem between a principal (a labor market authority) and an agent (a private firm) that is biased against applicants from a particular demographic group. The strength of the hiring regulation is limited by the fact that the firm has private information about the quality of its match with different job candidates. As in the delegation literature, we restrict attention to regulations that do not involve direct monetary transfers between the labor market authority and the firm. We show that the optimal regulation is dynamic, in the sense that its prescriptions depend on the firm's history of hiring. After any history, the optimal regulation either specifies that the firm is free to hire whichever candidate it sees fit (delegation state) or that the hiring process is directly controlled by the labor market authority (control state). In the delegation state, the firm is rewarded for hiring applicants from the discriminated group, and it is punished for passing on the opportunity to hire applicants from the discriminated group. The rewards and punishments give the firm an incentive to act less biased when making its hiring decision. The reward is delivered as an extension in the expected time until the hiring process is taken away from the firm, as well as a weakening of the incentives to hire discriminated applicants–which allows the firm to lean into its bias. Conversely, the punishment is delivered as a reduction in the expected time until the hiring process is taken away from the firm, as well as a strengthening of the incentives to hire discriminated applicants–which forces the firm to hire in a less biased way. In the control state, the labor market authority hires only applicants from the discriminated group. The control state is reached with probability 1 and it is absorbing.

In Section 2, we consider a basic version of the delegation problem described above. In every period, the agent has the option to hire an applicant from a particular demographic group. If the applicant is hired, the agent's payoff is given by $x - \eta$, where x denotes the quality of the applicant, and η denotes the extent of the agent's bias against the applicant's demographic group. The payoff to the principal is x. If the applicant is not hired, the payoff to the agent and the payoff to the principal are normalized to 0. Note that the principal is benevolent and unbiased, in the sense that it cares about the quality of the applicant hired by the agent, but does not care about the bias of the agent. The quality of the applicant x is a continuous random variable drawn from a distribution F, and its realization is privately observed by the agent.

In the first period, the principal commits to a direct-revelation mechanism without transfers in order to maximize its own value, which would be equal to the value received by the agent if the agent were unbiased. In each period, the mechanism solicits a report from the agent about the quality of the applicant. Depending on the realized history of play, the mechanism may make use of the agent's report to decide whether the applicant is hired, or it may ignore the agent's report. In the first case, the mechanism produces the same outcomes as if the agent was directly choosing whether or not to hire the applicant, taking as given the consequences of its decision on the future prescriptions of the mechanism. For this reason, we refer to the first case as "delegation". In the second case, the mechanism produces the same outcomes as if the principal was directly choosing whether to hire the applicant. For this reason, we refer to the second case as "control". Since the upper envelope of the principal's value under delegation and control is unlikely to be a concave function of the agent's value, the mechanism includes a public lottery between delegation and control at the beginning of each period. In Section 2, we restrict attention to mechanisms with the property that: (i) once the mechanism gives control to the principal, the principal remains in control forever; (ii) while in control, the principal hires every applicant. In Section 3, we prove that these restrictions are made without any loss in generality.

We formulate the optimal mechanism design problem recursively, using the agent's promised value as an auxiliary state variable. We first characterize the optimal lottery between delegation and control as a function of the agent's promised value V. We show that, if V is lower than some threshold V_C, the optimal lottery is non-degenerate, in the sense that it assigns positive probability to both delegation and control. If the outcome of the lottery is control, the principal hires every future applicant and the agent's value falls to its lower bound V_P. If the outcome of the lottery is delegation, the agent keeps making the hiring decisions and its value moves up to the threshold V_C. If the agent's promised value V is higher than V_C, the optimal lottery is degenerate. In particular, the hiring decision is delegated to the agent and the agent's continuation value $\hat{V}$ is equal to its promised value V.

We then characterize the optimal incentives in delegation as a function of the agent's promised value $\hat{V}$. We show that, if the agent hires the applicant in the current period, it gets rewarded with a continuation value V_1 that is strictly greater than the promised value $\hat{V}$. If the agent does not hire the applicant, it gets punished with a continuation value V_0 that is strictly smaller than the promised value $\hat{V}$. The gap between V_1 and V_0 induces the agent's to adopt a reservation quality for hiring, R, that is tilted away from η, the agent's preferred reservation quality, and towards 0, the principal's preferred reservation quality.

The mechanism delivers the reward $V_1 - \hat{V}$ by increasing the expected time until the hiring decision is taken away from the agent, and by weakening the incentives faced by the agent, which allows the agent to adopt a reservation quality R that is closer to its preferred one. Conversely, the mechanism delivers the punishment $\hat{V} - V_0$ by decreasing the expected time until the hiring decision is taken away from the agent, and by strengthening the incentives faced by the agent, which forces the agent to adopt a reservation quality R that is further away from its preferred one. If the agent keeps hiring applicants, its value approaches V_F, which denotes the value that the agent could obtain if it had permanent discretion over hiring. In this limit, the hiring incentives provided by the mechanism vanish, and the agent's reservation quality R approaches η. If the agent keeps on passing up on applicants, its value eventually falls below the threshold V_C, and the agent faces the threat of control, where applicants are hired irrespective of their quality. Even though the probability that the agent hires the applicant is always strictly positive, the mechanism reaches the control state with probability 1.

In Section 4, we establish that the characterization of the optimal mechanism in the baseline environment can be directly applied to richer, and more realistic situations. First, we show that the optimal mechanism in the baseline environment is identical to the optimal mechanism in a model where the agent can choose between an applicant from a demographic group against which it is biased (a contentious applicant) and an applicant from a demographic group against which it holds no bias (an uncontentious applicant). Second, we show that the analysis of the optimal mechanism in the baseline environment can be extended to the case in which the agent can choose between n contentious applicants and m uncontentious applicants. Third, we show that the analysis of the baseline environment can be applied to the case in which the agent's bias is positive rather than negative. Lastly, we show that the analysis can be generalized to the case in which the agent privately observes whether a contentious applicant is available for hire or not.

Our paper relates to the literature on delegation, broadly defined as mechanism design without transfers (Holmstrom 1977, Alonso and Matouschek 2008). Important examples of delegation in static environments include allocating decision rights within a firm (Aghion and Tirole 1997), setting hiring rules for a biased employer (Frankel 2021), and setting trade tariffs (Amador and Bagwell 2013). Several papers study delegation in dynamic models with hyperbolically-discounting agents, and focus on designing rules for time-inconsistent individuals (Angeletos, Werning and Amador 2006), governments (Halac and Yared 2018), or monetary authorities (Athey, Atkeson and Kehoe 2005). A common result in these papers is that, as long as shocks are independently drawn over time, the optimal mechanism is static—in the sense that it makes the same prescriptions independently of the realized history of play. In contrast to these papers, we find that the optimal mechanism is history-dependent. Intuitively this is so because in our model the principal and the agent disagree not only today, but also in the future.

Other papers on dynamic delegation find that the optimal mechanism is history-dependent. Jackson and Sonneschein (2007) show that "quota mechanisms" are optimal as players become infinitely patient. Escobar and Toikka (2013) establish a similar result with Markov types. In contrast to these papers, we consider an environment where players are impatient and, hence, the timing, and not simply the frequency, with which the agent takes a particular action matters to the principal. Frankel (2016) and Malenko (2019) show that appropriately discounted quotas are optimal also in an environment in which players are impatient, as long as the agent's preferences are state-independent. In contrast to these papers, we consider an environment in which both the principal and the agent care about the quality of the applicant. In such an environment, quotas are excessively restrictive.

The closest papers to ours are Li, Matouschek and Powell (2017), Guo and Horner (2020), and Lipnowski and Ramos (2020). Li, Matouschek and Powell (2017) consider a dynamic version of Aghion and Tirole (1997) in which the presence of a principal's preferred project is privately observed by the agent, who is rewarded for adopting it, and punished for undertaking its own preferred project. Guo and Horner (2020) study the problem of a principal that needs to rely on a biased agent to assess the quality of an investment opportunity. Lipnowski and Ramos (2020) consider the same environment as Guo and Horner (2020), but assume that the principal lacks commitment. These papers restrict attention to environments in which the quality of the investment is a binary random variable, while we consider an environment in which the quality of the applicant is a continuous random variable. The difference is not purely technical, as it affects the properties of the optimal mechanism. First, with a continuum of qualities, the mechanism rewards and punishes the agent by inducing changes in the agent's reservation quality. With two qualities, this margin for rewards and punishments is inactive. Second, with a continuum of qualities, we find that the mechanism drives the agent to its lowest value with probability 1. With two qualities, the mechanism may drive the agent to either its lowest or its highest value (see also the repeated moral-hazard problems in Clementi and Hopenhayn 2006, Sannikov 2008, and Padro i Miguel and Yared 2012). Third, with a continuum of qualities, it is immediate to extend the analysis of the optimal mechanism to an environment with multiple investment opportunities, some of which are contentious and some of which are not. This extension captures the essence of labor market discrimination.

2 Baseline model

In this section, we consider a basic environment. In every period, the principal has the option to hire a new applicant from some particular demographic group. The principal does not have the expertise to assess the quality of the applicant, but the agent does. The agent, however, is biased against applicants from that demographic group. The principal commits to a direct-revelation mechanism in order to elicit the information of the agent while minimizing the consequences of the agent's bias. The mechanism does not allow for monetary transfers between the principal and the agent and, in this sense, the mechanism design problem is a delegation problem. In Section 2.1, we describe and interpret the environment. In Section 2.2, we formulate the mechanism design problem recursively as a two-stage problem, using the agent's value as an auxiliary state variable. In Section 2.3, we characterize the solution of the first-stage problem—a lottery between delegation and control, where delegation means that, effectively, the agent directly chooses whether to hire the applicant, and control means that, effectively, the principal directly chooses whether to hire the applicant. In Section 2.4, we characterize the solution of the second-stage problem—the design of the mechanism when the hiring decision is delegated to the agent.

2.1 Environment

In every period $t = 0, 1, 2, ...$, a new applicant is available for hire. The quality x of the applicant is drawn from a continuously differentiable cumulative distribution function $F(x)$ with mean 0 and support $X = [\underline{x}, \overline{x}]$.[1] The agent observes the quality of the applicant, but the principal does not. If the applicant is hired, the agent obtains a flow payoff of $(1-\beta)(x-\eta)$ and the principal obtains a flow payoff of $(1-\beta)x$, where $\beta \in (0,1)$ is the factor at which the principal and the agent discount future payoffs, and $\eta \in (0, \overline{x})$ is the agent's bias against hiring the applicant. If the applicant is not hired, both the agent and the principal obtain a flow payoff of 0. No monetary transfers between the principal and the agent are allowed.

In period $t = 0$, the principal commits to a direct-revelation mechanism. In every period t, the mechanism solicits a report $\hat{x} \in X$ from the agent about the quality $x \in X$ of the applicant and decides whether or not to hire the applicant $a \in \{0, 1\}$.[2] Depending on the history h_t of play observed up to period t, the mechanism may act on the report from the agent, in the sense that it does or does not hire the applicant depending on the report, or it may ignore the report from the agent, in the sense that it may or may not hire the applicant irrespective of the report. In the first case, we say that the mechanism is delegating the hiring decision to the agent. In the second case, we say that the mechanism is letting the principal control the hiring decision.

Consider the case in which hiring is delegated to the agent. Let X_0 denote the set

[1] The assumption $\mathbb{E}[x] = 0$ is not critical for any of our results. The cases in which $\mathbb{E}[x] < \eta$ only require minor but tedious modifications in the proofs. The cases in which $\mathbb{E}[x] \geq \eta$ are more involved. Numerically, however, we find that our results hold also when $\mathbb{E}[x] \geq \eta$.

[2] We restrict attention to mechanisms that specify a hiring probability $a(x)$ that is either 0 or 1. The restriction is not without loss in generality. The restriction, however, is natural. If the mechanism specifies a hiring probability $a(x) \in \{0, 1\}$, a direct-revelation mechanism can be implemented by delegating the hiring decision to the agent or letting the principal control the hiring decision, depending on the hirtory of play. In this sense, the direct-revelation mechanism is implemnted as a genuine delegation mechanism– a mechanism that assigns the decision power to either player depending on the history of play. If the mechanism is allowed to specify a hiring probability $a(x) \in [0, 1]$, the direct-revelation mechanism cannot be implemented as a delegation mechanism.

of reports such that the mechanism does not hire the applicant, i.e. $a(\hat{x}, h_t) = 0$ for all $\hat{x} \in X_0$. Similarly, let X_1 denote the set of reports such that the mechanism does hire the applicant, i.e. $a(\hat{x}, h_t) = 1$ for all $\hat{x} \in X_1$. If the agent reports $\hat{x}$, its flow payoff is $a(\hat{x}, h_t)(1-\beta)(x-\eta)$ and its continuation value is some $V(\hat{x}, h_t)$. Since a direct-revelation mechanism must induce the agent to report the applicant's quality truthfully, it follows that $V(\hat{x}, h_t) = V_0(h_t)$ for all $\hat{x} \in X_0$ and $V(\hat{x}, h_t) = V_1(h_t)$ for all $\hat{x} \in X_1$. In words, the mechanism must give the same continuation value $V_0(h_t)$ to the agent for all the reports that make the mechanism pass on the applicant, and the same continuation value $V_1(h_t)$ for all the reports that make the mechanism hire the applicant. Moreover, since the mechanism must induce the agent to report the applicant's quality truthfully, it has to be the case that $X_0 = [x_\ell, R(h_t))$ and $X_1 = [R(h_t), x_h]$, where $R(h_t)$ is the quality of the applicant that makes the agent indifferent between $(1-\beta)(x-\eta) + \beta V_1(h_t)$ and $\beta V_0(h_t)$. Overall, when the mechanism uses the agent's report, it is as if the agent was directly choosing whether or not to hire the applicant, taking into account that its continuation value is $V_0(h_t)$ if it does not hire the applicant and $V_1(h_t)$ if it does.

Consider the case in which hiring is controlled by the principal. In this case, the mechanism either does not hire the applicant irrespective of the agent's report, i.e. $a(\hat{x}, h_t) = 0$ for all $\hat{x} \in X$, or it hires the applicant irrespective of the agent's report, i.e. $a(\hat{x}, h_t) = 1$ for all $\hat{x} \in X$. In either case, since the mechanism must induce the agent to report truthfully the quality of the applicant, the agent's continuation value $V(\hat{x}, h_t)$ must be independent of the agent's report $\hat{x}$. Overall, when the mechanism is ignoring the agent's report, it is as if the principal was controlling the hiring process directly.

Clearly, the combinations of agent's and principal's values that the mechanism can implement under delegation and under control are different and, in turn, the upper envelope of the agent's and principal's values under delegation and under control need not be concave. For this reason, it is natural to let the mechanism specify a public lottery between delegation and control at the beginning of every period t. Specifically, at the beginning of every period, the mechanism specifies some probability $p(h_t)$ with which hiring is controlled by the principal and some probability $1-p(h_t)$ with which hiring is delegated to the agent.

In this section, we are going to restrict attention to mechanisms such that: (i) if the mechanism hands the control of the hiring process to the principal, it does so forever; (ii) if the mechanism hands control of the hiring process to the principal, the mechanism instructs the principal to hire every applicant. In the next section, we are going to show that these restrictions are without loss in generality. Even though starting with a restriction on the space of mechanisms and then showing that it is without loss in generality may appear awkward to some of our readers, it does lead to a neater derivation of the optimal mechanism.

Before turning to the characterization of the optimal mechanism, let us briefly comment on the environment. At an abstract level, the environment is typical in the dynamic delegation literature. In each period, the principal has to decide whether or not to invest in a particular project, but some of the information that is relevant for making the decision is privately held by an agent whose preferences are only partially aligned with those of the principal. The principal commits to a direct-revelation mechanism without monetary transfers in order to make use of the agent's private information while mitigating the consequences of the agent's bias. The key difference between our environment and those studied by the existing dynamic delegation literature is that the quality x of the investment is not a binary random variable, but a continuous random variable with an arbitrary distribution F. The continuity of the random variable x is not a simple technicality. Indeed, the continuity of the random variable x allows us to immediately generalize the characterization of the optimal mechanism to richer and more realistic cases. For instance, the characterization of the optimal mechanism generalizes to the case in which the principal can invest in one of n available projects or not invest at all. The characterization of the optimal mechanism also generalizes to the case in which the principal can invest in one of several projects that come in two observable types, one type against which the agent is biased and one type for which the agent is neutral.

Concretely, our preferred interpretation of the environment is about discrimination in the labor market. In our preferred interpretation, the principal is a labor market authority and the agent is a firm. Through some search-and-matching process, an applicant from a particular demographic group becomes available to the firm for hiring. The firm and the labor market authority both care about the quality of the applicant, but only the firm is in the position to evaluate the applicant's quality. The firm and the labor market authority disagree about the value of hiring applicants from that the demographic group. Specifically, the labor market authority values these hires more than the firm. The disagreement may arise directly from the fact that the firm is biased against the applicant's demographic group, while the labor market authority is not. The disagreement may obtain indirectly from the fact that the labor market authority values the long-term benefits of a diverse workforce, while the firm does not. The labor market authority commits to a hiring regulation in order to make use of the firm's expertise about applicants while mitigating the consequences of the firm's bias.

In the environment described above, an applicant from the discriminated demographic group does not compete against any other current or future applicants because, say, the firm operates a technology with constant returns to scale in labor. The analysis of the basic environment will be extended to more realistic cases in which several applicants compete for the same job because, say, the firm operates a technology with decreasing returns to scale. In one of these cases, all applicants come from the discriminated demographic group. In another case, some applicants come from a demographic group against which the firm is biased, while other applicants come from a group against which the firm holds no bias.

2.2 Recursive formulation and preliminaries

The principal's mechanism design problem can be formulated recursively, using the value of the mechanism to the agent as an auxiliary state variable. In the first-stage of the recursive problem, the mechanism chooses a lottery between delegation and control, subject to delivering a particular value to the agent. In the second-stage, which is the stage associated with the delegation branch of the lottery, the mechanism chooses the agent's continuation value conditional on the agent hiring the applicant, and the agent's continuation value conditional on the agent not hiring the applicant, subject to delivering a particular expected value to the agent.

Formally, the first-stage problem is

$$J(V) = \max_{p \in [0,1], \hat{V} \in \hat{\mathcal{V}}} pJ_P + (1-p)\hat{J}(\hat{V}), \tag{2.1}$$

subject to the promise-keeping constraint

$$V = pV_P + (1-p)\hat{V}. \tag{2.2}$$

The first-stage problem is easy to understand. The mechanism chooses the probability p with which it permanently assigns the hiring decision to the principal (control), the probability $1-p$ with which this period's hiring decision is made by the agent (delegation), as well as the value $\hat{V}$ to the agent in case the outcome of the lottery is delegation. The mechanism makes these choices to maximize the principal's expected value, subject to delivering some expected value of V to the agent. The value to the principal is J_P if hiring is controlled by the principal, and $\hat{J}(\hat{V})$ if hiring is delegated to the agent. The value to the agent is V_P if hiring is controlled by the principal, and $\hat{V}$ if hiring is delegated to the agent. The values J_P and V_P, with P standing for "punishment", are

$$J_P \equiv \int_{\underline{x}}^{\overline{x}} x dF(x) = 0,\ V_P \equiv \int_{\underline{x}}^{\overline{x}} (x - \eta)\, dF(x) = -\eta. \tag{2.3}$$

The second-stage problem is

$$\hat{J}(\hat{V}) = \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R x dF(x) + \beta \left[F(R) J(V_0) + (1 - F(R)) J(V_1) \right], \tag{2.4}$$

subject to the promise-keeping constraint

$$\hat{V} = (1-\beta) \int_R (x - \eta) dF(x) + \beta \left[F(R) V_0 + (1 - F(R))) V_1 \right], \tag{2.5}$$

and the incentive-compatibility constraint

$$R = \eta - \frac{\beta}{1-\beta} \left(V_1 - V_0 \right). \tag{2.6}$$

The second-stage problem is also easy to understand. The mechanism chooses the agent's continuation value V_0 conditional on the agent not hiring the applicant, and the agent's continuation value V_1 conditional on the the agent hiring the applicant. Given V_0 and V_1, the agent finds it optimal to hire the applicant if and only if the applicant's quality x exceeds the reservation threshold R in (2.6). The mechanism maximizes the principal's value, subject to delivering the expected value $\hat{V}$ to the agent. In the current period, the payoff to the principal is $(1-\beta)\int_R x dF(x)$ and the payoff to the agent is $(1-\beta)\int_R (x-\eta)dF(x)$. If the agent does not hire the applicant, an event that occurs with probability $F(R)$, the continuation value to the principal is $J(V_0)$ and the continuation value to the agent is V_0. If the agent hires the applicant, an event that occurs with probability $1-F(R)$, the continuation value to the principal is $J(V_1)$ and the continuation value to the agent is V_1.

The formulation of the first and second-stage problems (2.1) and (2.4) is still incomplete, as it does not specify the choice sets $\mathcal{V}$ and $\hat{\mathcal{V}}$ for the agent's continuation values. The agent's continuation value $\hat{V}$ in (2.1) must be implementable in the second stage, in the sense that there exists a mechanism that delivers $\hat{V}$ to the agent in the second stage. The agent's continuation values V_0 and V_1 in (2.4) must be implementable in the first stage, in the sense that there exist mechanisms that deliver V_0 and V_1 to the agent in the first stage. The set $\mathcal{V}$ denotes the agent's values that can be implemented in the first-stage. The set $\hat{\mathcal{V}}$ denotes the agent's values that can be implemented in the second-stage.

Lemma 1 below characterizes the implementable sets $\mathcal{V}$ and $\hat{\mathcal{V}}$. The lemma shows that, in the first stage, the implementable set is the interval $[V_P, V_F]$, where V_P is the value to the agent of a mechanism in which the principal has permanent control over hiring, and V_F is the value to the agent of a mechanism in which hiring is permanently delegated to the agent, with F standing for "freedom." The values V_P and V_F are implementable. Values between V_P and V_F can be implemented by choosing the appropriate lottery between control and delegation. Values lower than V_P and values greater than V_F cannot be implemented in the first stage. In the second stage, the implementable set is the interval $[V_L, V_F]$, where V_L is the value to the agent of a mechanism in which hiring is in the hands of the agent in the current period and in the hands of the principal from the next period onwards. Values between V_L and V_F can be implemented by appropriately choosing the agents' continuation values. Values lower than V_L and greater than V_F cannot be implemented in the second stage.

Lemma 1: (Implementability) *The sets V and $\hat{V}$ are, respectively, given by*

$$\mathcal{V} = [V_P, V_F], \quad \hat{\mathcal{V}} = [V_L, V_F], \tag{2.7}$$

where V_L and V_F are defined as

$$V_L \equiv (1-\beta)V_F + \beta V_P, \tag{2.8}$$

$$V_F \equiv \int_{\underline{x}}^{\overline{x}} \max\{x-\eta, 0\} dF(x) = \int_{\eta} (x-\eta) dF(x). \tag{2.9}$$

Proof: Consider the first-stage problem (2.1). The choice $p = 1$ is feasible. For $p = 1$, the first-stage value to the agent is V_P. Hence, $V_P \in \mathcal{V}$. Consider the second-stage problem (2.4). The choices $V_0 = V_P$ and $V_1 = V_P$ are feasible. Given $(V_0, V_1) = (V_P, V_P)$, the second-stage value to the agent is

$$\begin{aligned} \overline{V}_1 &= (1-\beta)\int_{\eta}(x-\eta)dF(x) + \beta V_P \\ &= (1-\beta)V_F + \beta V_P, \end{aligned} \tag{2.10}$$

where the first line in (2.10) makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the second line in (2.10) makes use of the definition of V_F. Hence, $\overline{V}_1 \in \hat{\mathcal{V}}$. Notice that, since $V_F > V_P$, $\overline{V}_1 > V_P$ and $\overline{V}_1 < V_F$.

Return to the first-stage problem (2.1). The choices $p = (\overline{V}_1 - V)/(\overline{V}_1 - V_P)$ and $\hat{V} = \overline{V}_1$ are feasible for any $V \in [V_P, \overline{V}_1]$, since $p \in [0,1]$ and $\overline{V}_1 \in \hat{\mathcal{V}}$. Given the choices $p = (\overline{V}_1 - V)/(\overline{V}_1 - V_P)$ and $\hat{V} = \overline{V}_1$, the first-stage value to the agent is V. Hence, any $V \in [V_P, \overline{V}_1]$ belongs to $\mathcal{V}$. Now, consider the second-stage problem (2.4). The choices $V_0 = V$ and $V_1 = V$ are feasible for any $V \in [V_P, \overline{V}_1]$. Given $(V_0, V_1) = (V, V)$, the second-stage value to the agent is $\hat{V} = (1-\beta)V_F + \beta V$. Hence, any value $\hat{V}$ in the interval $[\overline{V}_1, \overline{V}_2]$ belongs to $\hat{\mathcal{V}}$, where

$$\overline{V}_2 = (1-\beta)V_F + \beta \overline{V}_1. \tag{2.11}$$

Since $\overline{V}_1 < V_F$, it follows that $\overline{V}_2 > \overline{V}_1$ and $\overline{V}_2 < V_F$.

Repeating the above argument k times yields that any value $V \in [V_P, \overline{V}_k]$ belongs to $\mathcal{V}$, and any value $\hat{V} \in [\overline{V}_1, \overline{V}_{k+1}]$ belongs to $\hat{\mathcal{V}}$, where

$$\overline{V}_{k+1} = (1-\beta)V_F + \beta \overline{V}_k. \tag{2.12}$$

The sequence $\{\overline{V}_i\}_{i=1}^{k+1}$ is strictly increasing and converges to V_F. Since a mechanism that permanently delegates hiring to the agent is worth V_F to the agent, V_F belongs to both $\mathcal{V}$ and $\hat{\mathcal{V}}$. These observations imply that $[V_P, V_F] \subseteq \mathcal{V}$ and $[\overline{V}_1, V_F] \subseteq \hat{\mathcal{V}}$.

Now, notice that the periodical payoff to the agent is such that

$$v \geq \min\left\{\min_R (1-\beta)\int_R (x-\eta)dF(x), (1-\beta)\int_{\underline{x}}^{\overline{x}} (x-\eta)dF(x)\right\}. \tag{2.13}$$

The right-hand side of (2.13) is the minimum between two terms. The first term is the lowest periodical payoff that an agent can attain when hiring is delegated to him. The second term is the periodical payoff that an agent attains when hiring is controlled by the principal. The minimum between the two terms is $-(1-\beta)\eta = (1-\beta)V_P$. This observation implies that a mechanism cannot implement any agent's value V smaller than V_P. That is, $V \notin \mathcal{V}, \hat{\mathcal{V}}$ for any $V < V_P$.

The periodical payoff v to the agent is such that

$$v \leq \max\left\{\max_R (1-\beta)\int_R (x-\eta)dF(x), (1-\beta)\int_{\underline{x}}^{\overline{x}}(x-\eta)dF(x)\right\}. \tag{2.14}$$

The right-hand side of (2.14) is the maximum between two terms. The first term is the highest periodical payoff that the agent can attain when hiring is delegated to him. The second term is the periodical payoff that the agent attains when the hiring is controlled by the principal. The maximum of the two terms is $(1-\beta)V_F$. Therefore, a mechanism cannot implement any value V strictly greater than V_F. That is, $V \notin \mathcal{V}, \hat{\mathcal{V}}$ for any $V > V_F$.

Finally, notice that $V \notin \hat{\mathcal{V}}$ for any $V < \overline{V}_1 = V_L$. To see why this is the case notice that, in the second stage, the agent's value is such that

$$\begin{aligned}
\hat{V} &= \max_R (1-\beta)\textstyle\int_R (x-\eta)dF(x) + \beta F(R)V_0 + \beta(1-F(R))V_1 \\
&\geq \max_R (1-\beta)\textstyle\int_R (x-\eta)dF(x) + \beta F(R)V_P + \beta(1-F(R))V_P \\
&= (1-\beta)V_F + \beta V_P = V_L,
\end{aligned} \tag{2.15}$$

where the second line makes use of the fact that V_0 and V_1 must be greater than V_P, and the third line makes use of the fact that the agent finds it optimal to set $R = \eta$ when V_0 and V_1 are both equal to V_P. ■

Before turning to the characterization of the optimal mechanism, it is useful to examine some points along the (V, J) and $(\hat{V}, \hat{J})$ frontiers. First, consider the agent's value V_P. Since $V_P \in \mathcal{V}$ and $V_P \notin \hat{\mathcal{V}}$, a mechanism can deliver the value V_P to the agent only in the first stage. It follows immediately from (2.1) that a mechanism can only deliver V_P to the agent through a lottery such that the hiring decision is delegated to the agent with probability 0 and permanently controlled by the principal with probability 1. Hence, $J(V_P) = J_P$. We denote as P the point (V_P, J_P). The point P belongs to the (V, J) frontier.

Second, consider the agent's value V_F. Since $V_F \in \mathcal{V}$ and $V_F \in \hat{\mathcal{V}}$, a mechanism can deliver the value V_F to the agent in both the first and the second stage. In the first stage, the mechanism can deliver the value V_F to the agent only through a lottery such that the hiring decision is delegated to the agent with probability 1 and gives the agent a second-stage value of V_F. Hence, $J(V_F)$ is equal to $\hat{J}(V_F)$. In the second stage, the mechanism can only deliver the value V_F to the agent by setting both V_0 and V_1 equal to V_F. Since $V_0 = V_1$ implies $R = \eta$, it follows that $\hat{J}(V_F)$ is equal to $(1-\beta)\int_\eta x dF(x) + \beta J(V_F)$. Since $J(V_F) = \hat{J}(V_F)$, it follows that $\hat{J}(V_F)$ is equal to $\int_\eta x dF(x)$, which we denote as J_F. We

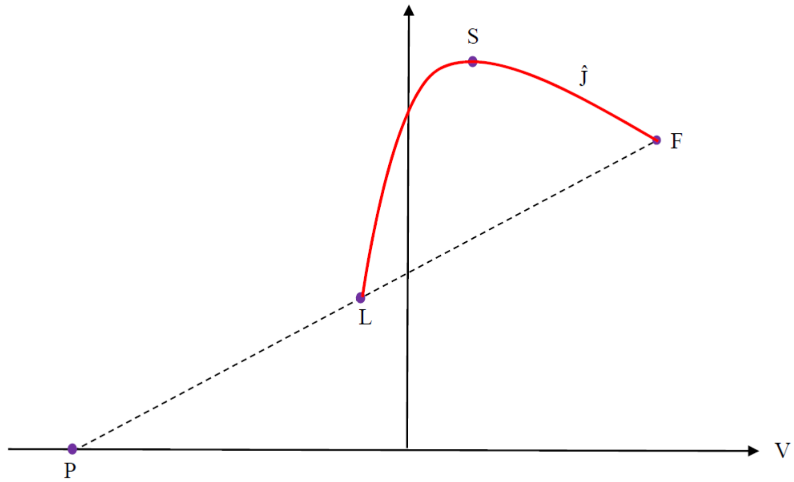


Figure 1: Points P, L, F and S, and frontier $(\hat{V}, \hat{J}(\hat{V}))$.

denote as F the point (V_F, J_F), which belongs to both the (V, J) and $(\hat{V}, \hat{J})$ frontiers. Note that, from the above characterization of the optimal mechanism, it follows that F is an absorbing state, in the sense that, once the mechanism reaches F, it remains there forever and, hence, the hiring decision is permanently delegated to the agent.

Third, consider the agent's value V_L, which is defined as $(1-\beta)V_F + \beta V_P$. Since V_L belongs to $\hat{\mathcal{V}}$, a mechanism can deliver the value V_L to the agent in the second stage. In the second stage, the mechanism can only deliver the value V_L to the agent by setting both V_0 and V_1 to V_P. Since $V_0 = V_1$ implies $R = \eta$, $\hat{J}(V_L)$ is equal to $(1-\beta)\int_\eta x dF(x) + \beta J(V_P)$. Since $J(V_P) = J_P$ and $\int_\eta x dF(x) = J_F$, it follows that $\hat{J}(V_L)$ is equal to $(1-\beta)J_F + \beta J_P$, which we denote as J_L. We denote as L the point (V_L, J_L). The point L belongs to the $(\hat{V}, \hat{J})$ frontier but, as we will show soon, it does not belong to the (V, J) frontier.

Note that the points P, L and F are on a line, as illustrated in Figure 1. Formally, P, L and F are on the line

$$H(V) = J_P + \frac{J_F - J_P}{V_F - V_P}(V - V_P). \tag{2.16}$$

The slope of $H(V)$ is

$$H' = \frac{J_F - J_P}{V_F - V_P}. \tag{2.17}$$

The slope is strictly positive because $J_F = \int_\eta x dF(x) > 0$, $J_P = \int_{\underline{x}}^{\overline{x}} x dF(x) = 0$, $V_F = \int_\eta (x-\eta) dF(x) > 0$ and $V_P = \int_{\underline{x}}^{\overline{x}} (x-\eta) dF(x) = -\eta$. The slope is smaller than 1 because $V_F - V_P$ is equal to $J_F - J_P + \eta F(\eta)$.

Lastly, we denote as S the point (V^*, J^*), where V^* is the agent's value such that the principal's value $J(V)$ is maximized. That is, V^* and J^* are, respectively, the arg-maximum and the maximum of $J(V)$ with respect to $V \in \mathcal{V}$. The point S is the initial position of the optimal mechanism, since the initial value of the mechanism to the agent is unconstrained by prior promises. For some parameter values, S is equal to F. In these cases, the optimal mechanism is such that the hiring decision is permanently delegated to the agent. For other parameter values, S is different from F and, hence, $J^* > J_F$ and $V^* < V_F$. In these cases, the optimal mechanism must be such that, after some history, the principal takes over the hiring process. In any case, V^* and J^* are also the arg-maximum and the maximum of $\hat{J}(\hat{V})$ with respect to $\hat{V} \in \hat{\mathcal{V}}$. This is so because $J(V)$ is a convex combination between J_P and some $\hat{J}(\hat{V})$, and $\hat{J}(V_F) > J_P$. Therefore, the optimal mechanism is always such that the hiring decision is initially delegated to the agent.

The following lemma identifies a sufficient condition on the parameters of the model such that $S \neq F$ and, hence, the optimal mechanism is dynamic. In the remainder of the paper, we assume that the sufficient condition holds.

Lemma 2: (Sufficient condition for a dynamic mechanism) *The starting position $S = (V^*, J^*)$ of the optimal mechanism is such that $J^* > J_F$ as long as η and F are such that*

$$\frac{\eta F'(\eta)}{F(\eta)} > \frac{\int_\eta x dF(x)}{\int_\eta x dF(x) + \eta F(\eta)}. \tag{2.18}$$

Proof: Consider the following mechanism. In the current period, the hiring decision is delegated to the agent. If the agent hires the applicant, the agent's continuation value V_1 is V_F. That is, if the agent hires the applicant, the hiring decision is permanently delegated to the agent. If the agent does not hire the applicant, the agent's continuation value V_0 is $\epsilon V_P + (1-\epsilon)V_F$ for some $\epsilon > 0$. The continuation value V_0 is delivered as a lottery that assigns probability ϵ to V_P and probability $1-\epsilon$ to V_F. That is, if the agent does not hire the applicant, the hiring decision is permanently delegated to the agent with probability $1-\epsilon$, and it is permanently controlled by the principal with probability ϵ. The mechanism need not be optimal, but it is feasible.

The value $\Gamma(\epsilon)$ of the mechanism to the principal is

$$\Gamma(\epsilon) = (1-\beta)\int_R xF'(x)dx + \beta F(R)\left[\epsilon J_P + (1-\epsilon)J_F\right] + \beta(1-F(R))J_F, \tag{2.19}$$

where the reservation quality R is

$$R = \eta - \frac{\beta}{1-\beta}\left[V_F - \epsilon V_P - (1-\epsilon)V_F\right]. \tag{2.20}$$

The derivative of $\Gamma(\epsilon)$ with respect to ϵ is

$$\Gamma'(\epsilon) = -\beta F(R)\left(J_F - J_P\right) - \frac{dR}{d\epsilon}\left\{(1-\beta)RF'(R) + \beta F'(R)\epsilon\left[J_F - J_P\right]\right\}, \tag{2.21}$$

where

$$\frac{dR}{d\epsilon} = -\frac{\beta}{1-\beta}(V_F - V_P). \tag{2.22}$$

When evaluated at $\epsilon = 0$, the derivative of $\Gamma(\epsilon)$ with respect to ϵ becomes

$$\Gamma'(0) = -\beta F(\eta)\,(J_F - J_P) + \beta\eta F'(\eta)(V_F - V_P). \tag{2.23}$$

The above expression is strictly positive as long as

$$\frac{\eta F'(\eta)}{F(\eta)} > \frac{J_F - J_P}{V_F - V_P} = \frac{\int_\eta x dF(x)}{\int_\eta x dF(x) + \eta F(\eta)}. \tag{2.24}$$

Note that the value $\Gamma(\epsilon)$ is such that $\Gamma(0) = J_F$. Hence, if condition (2.24) is satisfied, there exists some $\epsilon^* > 0$ such that $\Gamma(\epsilon^*) > J_F$. Next, note that the value $\Gamma(\epsilon^*)$ is such that $\Gamma(\epsilon^*) \leq J^*$, since the mechanism is feasible but need not be optimal. Hence, if condition (2.24) is satisfied, $J^* > J_F$. ■

Intuitively, condition (2.18) is a lower bound on the agent's bias η given the distribution F of applicant's quality. For instance, if the distribution of applicant's quality is uniform over some interval $[-\delta, \delta]$, condition (2.18) boils down to $\eta > \delta/2$. That is, as long as the agent's bias is large relative to the dispersion of the applicant's quality, the optimal mechanism involves a transfer of the hiring decision from the agent to the principal.[3]

2.3 Optimal lottery

In this subsection, we characterize the solution to the first-stage problem in (2.1), which is the design of the lottery between delegation and control. In order to simplify the characterization of the problem, we assume that the second-stage value to the principal, $\hat{J}(\hat{V})$, is strictly concave and differentiable. In order to guarantee the weak concavity of $\hat{J}(\hat{V})$, it would be enough to allow for lotteries in the second-stage problem (2.4). Doing so, however, would make the notation quite cumbersome. Weak concavity of $\hat{J}(\hat{V})$ would be enough to derive results that are analogous to those presented below. It is harder to find conditions to guarantee the differentiability of $\hat{J}(\hat{V})$. This is because, in contrast to the standard dynamic problems considered in, e.g., Stokey, Lucas and Prescott (1988), the only way to deliver a higher/lower value to the agent is by changing its continuation values.[4]

[3] Even in the case of other standard distributions, such as a Normal or an Exponential, condition (2.24) implies a lower bound on η that is increasing in the dispersion of the applicant's quality.

[4] Using an envelope argument, we can show that $\hat{J}(\hat{V})$ is differentiable everywhere if, for every $\hat{V}$, there is a history that is reached with positive probability where the first-stage lottery is non-degenerate. In our numerical examples, we find that this is indeed the case and, hence, the assumption of a differentiable $\hat{J}(\hat{V})$ is vindicated.

Using the promise-keeping constraint (2.2) to substitute out p, we can rewrite (2.1) as

$$J(V) = \max_{\hat{V} \in \hat{\mathcal{V}}} \frac{\hat{V} - V}{\hat{V} - V_P} J_P + \frac{V - V_P}{\hat{V} - V_P} \hat{J}(\hat{V}), \quad \text{s.t. } \hat{V} \geq V. \tag{2.25}$$

The necessary condition for the optimality of $\hat{V}$ is

$$\frac{V - V_P}{\hat{V} - V_P} \left[\hat{J}'(\hat{V}) - \frac{\hat{J}(\hat{V}) - J_P}{\hat{V} - V_P} \right] \begin{array}{ll} = 0 & \text{if } \hat{V} \in (\max\{V, V_\ell\}, V_F), \\ \leq 0, & \text{if } \hat{V} = \max\{V, V_\ell\}, \\ \geq 0, & \text{if } \hat{V} = V_F. \end{array} \tag{2.26}$$

Let us examine the term in square brackets on the left-hand side of (2.26). The function $\hat{J}'(\hat{V})$ is the derivative of the second-stage problem value function. Since $\hat{J}(\hat{V})$ is strictly concave and attains its maximum at some $V^* \in (V_L, V_F)$, it follows that $\hat{J}'(\hat{V})$ is strictly decreasing and such that $\hat{J}'(V_L) > 0$, $\hat{J}'(V^*) = 0$, and $\hat{J}'(V_F) < 0$. Since $\hat{J}(\hat{V})$ is strictly concave and such that $\hat{J}(V_L) = H(V_L)$, $\hat{J}(V_F) = H(V_F)$ and $\hat{J}(V^*) > H(V_F) > H(V^*)$, it follows that $\hat{J}'(V_L) > H'$ and $\hat{J}'(V_F) < 0 < H'$. For the same reason, $\hat{J}(\hat{V}) > H(\hat{V})$ for all $\hat{V} \in (V_L, V_F)$.

The function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is the slope of the line connecting the points (V_P, J_P) and $(\hat{V}, \hat{J}(\hat{V}))$. Since $\hat{J}(V_L) = H(V_L)$ and $\hat{J}(V_F) = H(V_F)$, it follows that $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is equal to H' for $\hat{V} = V_L$ and $\hat{V} = V_F$. Since $\hat{J}(\hat{V}) > H(\hat{V})$ for all $\hat{V} \in (V_L, V_F)$, $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly greater than H' for all $\hat{V} \in (V_L, V_F)$. These observations imply that the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly smaller than the function $\hat{J}'(\hat{V})$ at $\hat{V} = V_L$ and strictly greater than $\hat{J}'(\hat{V})$ at $\hat{V} = V_F$. Therefore, there must exist at least one $V_C \in (V_L, V_F)$ such that $(\hat{J}(V_C) - J_P)/(V_C - V_P)$ is equal to $\hat{J}'(V_C)$. Clearly, any V_C is strictly greater than V_L. Similarly, any V_C is strictly smaller than V^*, since the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly positive and $\hat{J}'(\hat{V})$ is non-positive for any $\hat{V} \in [V^*, V_F]$.

It is easy to verify that the derivative of the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ with respect to $\hat{V}$ equals zero if and only if $\hat{V} = V_C$, and that the function attains a local maximum at V_C. Therefore, V_C must be unique and the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must attain its global maximum at V_C. Moreover, since $\hat{J}'(\hat{V})$ is strictly decreasing and the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is equal to $\hat{J}'(\hat{V})$ and attains its maximum at V_C, it follows that $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must be strictly smaller than $\hat{J}'(\hat{V})$ for all $\hat{V} \in [V_L, V_C)$. Similarly, $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must be strictly greater than $\hat{J}'(\hat{V})$ for all $\hat{V} \in (V_C, V_F]$.

The above observations imply that, for any $V \in [V_P, V_C)$, the optimality condition (2.26) and the promise-keeping constraint (2.2) are satisfied if and only if

$$\hat{V} = V_C, \quad p = \frac{V_C - V}{V_C - V_P} > 0. \tag{2.27}$$

Plugging the optimal choice for $\hat{V}$ in (2.25) yields

$$J(V) = J_P + \frac{V - V_P}{V_C - V_P}(J_C - J_P), \tag{2.28}$$

where $J_C \equiv \hat{J}(V_C)$. Differentiating (2.28) with respect to V yields

$$J'(V) = \frac{J_C - J_P}{V_C - V_P} < \hat{J}'(V), \tag{2.29}$$

where the last inequality follows from the fact that $(J_C - J_P)/(V_C - V_P) = \hat{J}'(V_C)$ and $\hat{J}'(V) > \hat{J}'(V_C)$.

For any $V \in [V_C, V_F]$, the optimality condition (2.26) and the promise-keeping constraint (2.2) are satisfied if and only if

$$\hat{V} = V, \quad p = 0. \tag{2.30}$$

Plugging the optimal choice for $\hat{V}$ in (2.25) yields

$$J(V) = \hat{J}(V). \tag{2.31}$$

Differentiating (2.31) with respect to V yields

$$J'(V) = \hat{J}'(V). \tag{2.32}$$

The proposition below summarizes the characterization of the first-stage problem.

Proposition 1: (Optimal lottery) *The solution to the first-stage problem (2.1) is such that:*

1. *For $V \in [V_P, V_C)$, the hiring process is permanently controlled by the principal with probability p, and it is delegated to the agent with probability $1 - p$, where $p = (V_C - V)/(V_C - V_P)$. Conditional on delegation, the agent's value is V_C.*
2. *For $V \in [V_C, V_F]$, the hiring process is permanently controlled by the principal with probability 0, and it is delegated to the agent with probability 1. Conditional on delegation, the agent's value is V.*
3. *The value to the principal is*

$$J(V) = \begin{cases} J_P + \dfrac{V - V_P}{V_C - V_P}(V - V_P), & \text{if } V \in [V_P, V_C), \\ \hat{J}(V), & \text{if } V \in [V_C, V_F]. \end{cases} \tag{2.33}$$

4. *The point $C = (V_C, J_C)$ is such that $J_C = \hat{J}(V_C)$ and $V_C \in (V_L, V^*)$ is the unique solution to*

$$\hat{J}'(V_C) = \frac{J_C - J_P}{V_C - V_P}. \tag{2.34}$$

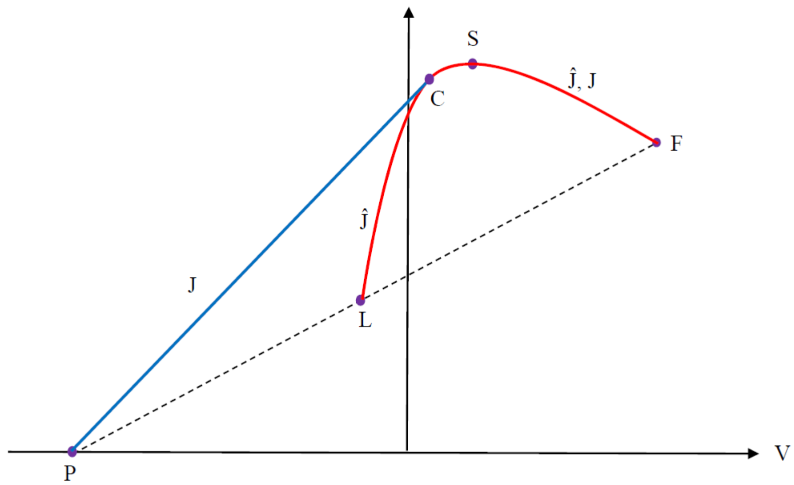


Figure 2: Value functions $J(V)$ and $\hat{J}(\hat{V})$.

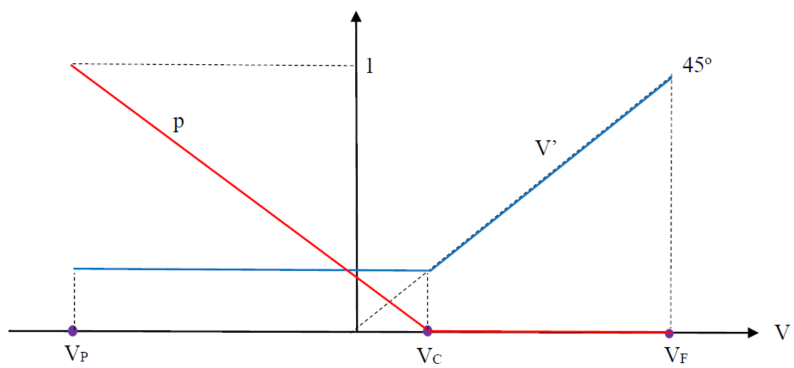


Figure 3: Optimal lottery p and $\hat{V}$ as function of the promised value V.

The properties of the optimal lottery between delegation and control are illustrated in Figures 2 and 3. The properties are intuitive. The lottery allows the mechanism to achieve any combinations of the agent's and the principal's values that lie on a line connecting the control values, given by the point P, and one of the delegation values, given by a point along the $(\hat{V}, \hat{J})$ frontier. The optimal lottery is the highest of such lines, which is the line that connects P and C, the point where the lottery line is tangent to the $(\hat{V}, \hat{J})$ frontier. It follows that, for any $V \in (V_P, V_C)$, the optimal lottery randomizes with between P and C with non-degenerate probabilities. For any $V \in [V_C, V_F]$, the lottery is useless, in the sense that the lottery assigns the hiring decision to the agent with probability 1.

2.4 Optimal incentives in delegation

In this subsection, we characterize the solution to the second-stage problem in (2.4), which is the design of the mechanism conditional on the hiring decision being delegated to the agent in the current period. We focus the analysis on agent's promised values $\hat{V}$ in the interval $[V_C, V_F]$, since Proposition 1 implies that values outside of this interval are never reached.

In Proposition 2 we show that the solution to the second-stage problem is such that the agent's reservation quality R is in the interior of the support of the quality distribution F. Proposition 2 is useful, as it guarantees that there is a strictly positive probability that the agent hires the applicant, in which case the agent's continuation value is V_1, and a strictly positive probability that the agent does not hire the applicant, in which case the agent's continuation value is V_0.

The intuition behind Proposition 2 is simple. If R was smaller than $\underline{x}$, the agent would hire the applicant for sure. The agent's and the principal's values would then be a non-degenerate convex combination between the point P and some point along the $(\hat{V}, \hat{J})$ frontier. This convex combination is achievable in the first-stage problem (2.1) but, as shown in Proposition 1, a non-degenerate convex combination is not optimal. Therefore, $\hat{J}(\hat{V})$ would be strictly smaller than $J(\hat{V})$. Proposition 1, however, guarantees that $J(\hat{V})$ is equal to $\hat{J}(\hat{V})$ for all $\hat{V} \in [V_C, V_F]$. Similarly, if R was greater than $\overline{x}$, the agent would not hire the applicant. Therefore, the agent's and principal's values would be a convex combination between the point $(0, 0)$ and some point on the frontier $(\hat{V}, \hat{J})$. We show that such a convex combination also implies that $\hat{J}(\hat{V})$ would be strictly smaller than $J(\hat{V})$, which would contradict Proposition 1.

Proposition 2: (Optimal reservation quality) *For all $\hat{V} \in [V_C, V_F]$, the solution to the second-stage problem (2.4) is such that the reservation quality R belongs to $(\underline{x}, \overline{x})$.*

Proof: On the way to a contradiction, suppose that, for some $\hat{V}^a \in [V_C, V_F]$, the solution to the second-stage problem (2.4) specifies continuation values V_0^a, $V_1^a \in \mathcal{V}$ such that

$$R^a = \eta - \frac{\beta}{1-\beta}(V_1^a - V_0^a) \leq \underline{x}. \tag{2.35}$$

In this case, the value to the agent is

$$\begin{aligned} \hat{V}^a &= (1-\beta)\int_{\underline{x}}^{\overline{x}}(x-\eta)dF(x) + \beta V_1^a \\ &= (1-\beta)V_P + \beta V_1^a. \end{aligned} \tag{2.36}$$

The value to the principal is

$$\begin{aligned} \hat{J}(\hat{V}^a) &= (1-\beta)\int_{\underline{x}}^{\overline{x}} x dF(x) + \beta J\left(V_1^a\right) \\ &= (1-\beta)J_P + \beta J(V_1^a) \\ &= (1-\beta)J_P + \beta \hat{J}(V_1^a) \end{aligned} \tag{2.37}$$

The second line in (2.36) makes use of the definition of V_P. The second line in (2.37) makes use of the definition of J_P. The third line in (2.37) makes use of the fact that $\hat{V}^a > V_P$ implies $V_1^a > \hat{V}^a \geq V_C$ and, in light of Proposition 1, $J(V_1^a) = \hat{J}(V_1^a)$.

Now, consider the first-stage problem (2.1) for $\hat{V}^a$. The problem is

$$\begin{aligned} J(\hat{V}^a) = \quad & \max_{p \in [0,1], \hat{V} \in \hat{\mathcal{V}}} pJ_P + (1-p)\hat{J}(\hat{V}), \\ & \text{s.t. } \hat{V}^a = pV_P + (1-p)\hat{V}. \end{aligned} \tag{2.38}$$

The lottery $(p, \hat{V}) = (\beta, V_1^a)$ is a feasible choice for (2.38), since $\beta \in (0,1)$, $V_1^a \in \hat{\mathcal{V}}$, and $\beta V_P + (1-\beta)V_1^a$ is equal to $\hat{V}^a$. The lottery (β, V_1^a) is not the optimal choice for (2.38), since $\hat{V}^a \geq V_C$ and Proposition 1 states that the optimal lottery for any $V \geq V_C$ is $(p, \hat{V}) = (0, V)$. From (2.37), it follows that the value to the principal of the feasible and suboptimal lottery (β, V_1^a) is $\hat{J}(\hat{V}^a)$. Since the lottery (β, V_1^a) is suboptimal, it follows that $\hat{J}(\hat{V}^a)$ is strictly smaller than $J(\hat{V}^a)$. However, Proposition 1 states that $J(V) = \hat{J}(V)$ for all $V \geq V_C$, including for $V = \hat{V}^a$. We have thus reached the desired contradiction.

On the way to a second contradiction, suppose that, for some $\hat{V}^b \in [V_C, V_F]$, the solution to the second-stage problem (2.4) specifies continuation values V_0^b, $V_1^b \in \mathcal{V}$ such that

$$R = \eta - \frac{\beta}{1-\beta}(V_1^b - V_0^b) \geq \overline{x}. \tag{2.39}$$

In this case, the value to the agent is

$$\hat{V}^b = \beta V_0^b. \tag{2.40}$$

The value to the principal is

$$\begin{aligned} \hat{J}(\hat{V}^b) \quad & = 0 + \beta J\left(V_0^b\right) \\ & < (1-\beta)J(0) + \beta J(V_0^b) \\ & \leq J((1-\beta)0 + \beta V_0^b) \\ & = J(\hat{V}^b) \end{aligned} \tag{2.41}$$

The second line in (2.41) makes use of the fact that $J(0) > 0$ because $J(V) \geq H(V) > 0$ for all $V > V_P$ and $V_P < 0$. The third line makes use of the fact that $J(V)$ is weakly concave. The fourth line makes use of the fact that $V_0^b = \hat{V}^b/\beta$. Comparing the first and the last line yields $\hat{J}(\hat{V}^b) < J(\hat{V}^b)$. However, $\hat{J}(\hat{V}^b) = J(\hat{V}^b)$, since Proposition 1 implies that $J(V) = \hat{J}(V)$ for all $V \geq V_C$, including for $V = \hat{V}^b$. We have thus reached the desired contradiction. ■

We now turn to the characterization of the optimal agent's continuation values in the second-stage problem. The necessary condition for the optimality of an interior V_0 is

$$\begin{aligned} 0 = \quad & \beta F(R) J'(V_0) + \beta F'(R) \left[J(V_0) - J(V_1) \right] \frac{dR}{dV_0} - (1-\beta) R F'(R) \frac{dR}{dV_0} \\ & + \lambda \left\{ \beta F(R) + \beta F'(R)(V_0 - V_1) \frac{dR}{dV_0} - (1-\beta)(R-\eta) F'(R) \frac{dR}{dV_0} \right\}, \end{aligned} \tag{2.42}$$

where λ denotes the Lagrange multiplier on the promise-keeping constraint (2.5) and dR/dV_0 denotes the derivative of the agent's reservation quality (2.6) with respect to V_0, i.e.

$$\frac{dR}{dV_0} = \frac{\beta}{1-\beta}. \tag{2.43}$$

The first line on the right-hand side of (2.42) is the derivative of the principal's value with respect to V_0. This derivative is given by the change in the principal's continuation value due to the change in V_0 and the change in the principal's flow payoff and continuation value due to change in the agent's reservation quality R. The second line on the right-hand side of (2.42) is the derivative of the agent's value with respect to V_0 multiplied by λ. This derivative is given by the change in the agent's continuation value due to the change in V_0 and the change in the agent's flow payoff and continuation value due to change in the agent's reservation quality R.

Using the fact that the agent's reservation quality R is given by (2.6) and that the derivative dR/dV_0 is given by (2.43), we can rewrite (2.42) as

$$J'(V_0) + \lambda = \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\}. \tag{2.44}$$

The necessary condition for optimality for an interior V_1 is

$$\begin{aligned} 0 = \quad & \beta (1 - F(R)) J'(V_1) + \beta F'(R) \left[J(V_0) - J(V_1) \right] \frac{dR}{dV_1} - (1-\beta) R F'(R) \frac{dR}{dV_1} \\ & + \lambda \left\{ \beta (1 - F(R)) + \beta F'(R)(V_0 - V_1) \frac{dR}{dV_1} - (1-\beta)(R-\eta) F'(R) \frac{dR}{dV_1} \right\}, \end{aligned} \tag{2.45}$$

where dR/dV_1 denotes the derivative of the agent's reservation quality (2.6) with respect to V_1, i.e.

$$\frac{dR}{dV_1} = -\frac{\beta}{1-\beta}. \tag{2.46}$$

Using the fact that the agent's reservation quality R is given by (2.6) and that the derivative dR/dV_1 is given by (2.46), we can rewrite (2.45) as

$$J'(V_1) + \lambda = -\frac{F'(R)}{1 - F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\}. \tag{2.47}$$

Finally, the derivative $\hat{J}'(\hat{V})$ of the principal's value with respect to the agent's promised value $\hat{V}$ is equal to the Lagrange multiplier λ on the promise-keeping constraint (2.5).

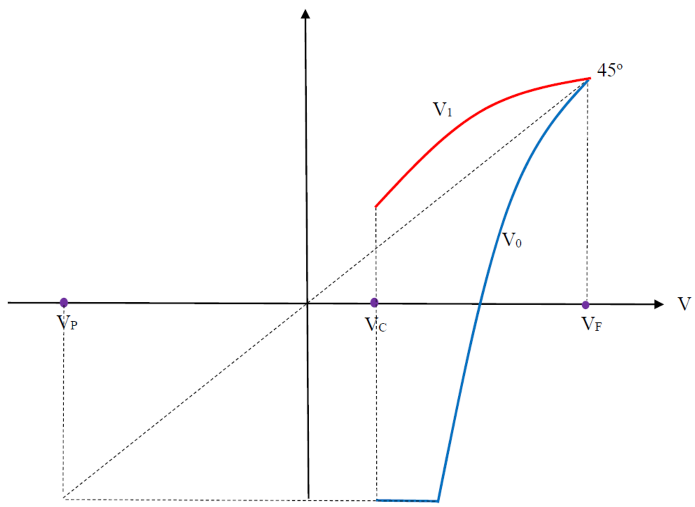


Figure 4: Optimal continuation values V_0 and V_1 as functions of the promised value $\hat{V}$.

That is,

$$\hat{J}'(\hat{V}) = -\lambda. \tag{2.48}$$

In the next proposition, we further characterize the solution to the second-stage problem (2.4). In order to simplify the analysis, let us take it as given that the agent's continuation values V_0 and V_1 are continuous functions of the value $\hat{V}$ promised to the agent. Using the first-order conditions (2.44) and (2.47) and the envelope condition (2.48), we show that, for any $\hat{V} \in (V_C, V_F)$, the agent's continuation value if it does not hire the applicant, V_0, is strictly smaller than the agent's promised value $\hat{V}$. Conversely, the agent's continuation value if it hires the applicant, V_1, is strictly greater than $\hat{V}$. Since $V_0 < V_1$, the agent's reservation quality R is strictly smaller than the agent's preferred quality cutoff η. Since $V_0 < \hat{V} < V_1$, the agent's value increases when it hires the applicant, and decreases when it does not hire the applicant. Similarly, for $\hat{V} = V_C$, we show that $V_0 \leq \hat{V} \leq V_1$ and $V_0 < V_1$ and, hence, $R < \eta$. As we already established, for $\hat{V} = V_F$, $V_0 = V_1 = V_F$ and, hence, $R = \eta$. We also show that V_1 is strictly smaller than V_F for all $\hat{V} \in [V_C, V_F)$, which implies that the optimal mechanism does never permanently delegate the hiring decision to the agent, independently of the agent's hiring record. The characterization results are illustrated in Figure 4.

Proposition 3: (Optimal incentives in delegation). *The solution to the second-stage problem (2.4) is such that:*

1. *For any agent's promised value $\hat{V} \in (V_C, V_F)$, the agent's continuation value conditional on not hiring the applicant is $V_0 < \hat{V}$, the agent's continuation value conditional on hiring the applicant is $V_1 > \hat{V}$, and the agent's reservation quality is $R < \eta$. For $\hat{V} = V_C$, $V_0 \leq V_C \leq V_1$, $V_0 < V_1$ and $R < \eta$. For $\hat{V} = V_F$, $V_0 = V_1 = V_F$ and $R = \eta$.*

2. *For any agent's promised value $\hat{V} \in [V_C, V_F)$, the agent's continuation value conditional on hiring the applicant is $V_1 < V_F$.*

Proof of part (1). We break down the proof in three claims.

Claim 1: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in [V_C, V_F)$, the continuation values V_0 and V_1 are different.

On the way to a contradiction, suppose that the continuation values V_0 and V_1 are both equal to some V, with $V \in (V_P, V_F)$. The necessary conditions (2.44) and (2.47) become

$$\begin{aligned} J'(V_0) + \lambda &= \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\} = \frac{F'(\eta)}{F(\eta)} \eta, \\ J'(V_1) + \lambda &= -\frac{F'(R)}{1-F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\} = -\frac{F'(\eta)}{1-F(\eta)} \eta, \end{aligned} \tag{2.49}$$

where the second equality in both lines makes use of the fact that $V_0 = V_1$ implies $J(V_1) = J(V_0)$ and $R = \eta$. The necessary conditions (2.49) imply that $J'(V_0) + \lambda > 0 > J'(V_1) + \lambda$. However, $V_0 = V_1$ implies that $J'(V_0) + \lambda = J'(V_1) + \lambda$. A contradiction.

On the way to another contradiction, suppose that the continuation values V_0 and V_1 are both equal to V_P. The promise-keeping constraint (2.5) becomes

$$\begin{aligned} \hat{V} &= (1-\beta) \int_R (x - \eta) dF(x) + \beta V_P \\ &= (1-\beta) \int_\eta (x - \eta) dF(x) + \beta V_P \\ &= (1-\beta) V_F + \beta V_P = V_L, \end{aligned} \tag{2.50}$$

where the second line makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the third line makes use of the definitions of V_F and V_L. Since $\hat{V} \geq V_C$ and $V_C > V_L$, we have reached a contradiction.

Lastly, suppose that the continuation values V_0 and V_1 are both equal to V_F. In this case, the promise-keeping constraint (2.5) becomes

$$\begin{aligned} \hat{V} &= (1-\beta) \int_R (x - \eta) dF(x) + \beta V_F \\ &= (1-\beta) \int_\eta (x - \eta) dF(x) + \beta V_F \\ &= (1-\beta) V_F + \beta V_F = V_F. \end{aligned} \tag{2.51}$$

where the second line makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the third lines makes use of the definition of V_F. Since $\hat{V} < V_F$, we have reached a contradiction.

Claim 2: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in [V_C, V_F)$, the continuation values V_0 and V_1 are such that $V_0 < V_1$.

Since $V_0 \neq V_1$ for all $\hat{V} \in [V_C, V_F)$ and V_0 and V_1 are continuous functions of $\hat{V}$, it follows that either $V_0 > V_1$ for all $\hat{V} \in [V_C, V_F)$ or $V_0 < V_1$ for all $\hat{V} \in [V_C, V_F)$. On the way to a contradiction, suppose that $V_0 > V_1$ for all $\hat{V} \in [V_C, V_F)$. If that is the case, the agent's reservation quality R is strictly greater than η for all $\hat{V} \in [V_C, V_F)$. For $\hat{V} = V_F$, the only feasible continuation values are $V_0 = V_1 = V_F$ and, hence, the agent's reservation quality R is equal to η. If hiring is delegated to the agent, $\hat{V} \in [V_C, V_F]$ and, hence, $R \geq \eta$. Therefore, if hiring is delegated to the agent, the principal's flow payoff is such that

$$\begin{aligned} u &\leq \max_{R \geq \eta} \int_R x dF(x) \\ &= \int_\eta x dF(x) = (1-\beta) J_F, \end{aligned} \tag{2.52}$$

where the second line uses the fact that the integral in the first line is strictly decreasing in R for all $R > 0$. If hiring is controlled by the principal, the principal's flow payoff is

$$u = \int_{\underline{x}}^{\overline{x}} x dF(x) = (1-\beta) J_P. \tag{2.53}$$

Since the value of the mechanism to the principal is the discounted sum of flow payoffs, it follows that

$$J(\hat{V}) \leq \frac{(1-\beta) \max\{J_F, J_P\}}{1-\beta} = J_F. \tag{2.54}$$

The above inequality holds for any $\hat{V} \in [V_L, V_F]$. However, $V^* \in [V_C, V_F)$ and $J(V^*) > J_F$. A contradiction.

Claim 3: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in (V_C, V_F)$, the continuation values V_0 and V_1 are such that $V_0 < \hat{V} < V_1$. By continuity, $V_0 \leq \hat{V} \leq V_1$ for $\hat{V} = V_C$.

Consider some value $\hat{V} \in (V_C, V_F)$ promised to the agent. The continuation values V_0 and V_1 are such that $V_0 < V_1$. Suppose that V_0 and V_1 are both interior. Using the envelope condition (2.48) to substitute out the Lagrange multiplier λ, we can rewrite the necessary conditions (2.44) and (2.47) as

$$\begin{aligned} J'(V_0) - \hat{J}'(\hat{V}) &= \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\}, \\ J'(V_1) - \hat{J}'(\hat{V}) &= -\frac{F'(R)}{1-F(R)} \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\}. \end{aligned} \tag{2.55}$$

Subtracting the second equation in (2.55) from the first yields

$$J'(V_0) - J'(V_1) = \left[\frac{F'(R)}{F(R)} + \frac{F'(R)}{1-F(R)} \right] \left\{ R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] \right\}. \tag{2.56}$$

Suppose that $V_0 \geq V_C$. Since $V_0 \geq V_C$ and $V_1 > V_0 \geq V_C$, Proposition 1 guarantees that $J'(V_0)$ is equal to $\hat{J}'(V_0)$ and $J'(V_1)$ is equal to $\hat{J}'(V_1)$. Since $V_1 > V_0$ and $\hat{J}$ is strictly concave, $\hat{J}'(V_0)$ is strictly greater than $\hat{J}'(V_1)$. Putting these observations together yields $J'(V_0) > J'(V_1)$. Since $J'(V_0) > J'(V_1)$, (2.56) implies

$$R + \frac{\beta}{1-\beta} \left[J(V_1) - J(V_0) \right] > 0. \tag{2.57}$$

Using (2.57) in the first necessary condition in (2.55) yields $J'(V_0) > \hat{J}'(\hat{V})$. Since $J'(V_0)$ is equal to $\hat{J}'(V_0)$ and $\hat{J}$ is strictly concave, $V_0 < \hat{V}$. Using (2.57) into the second necessary condition in (2.55) yields $J'(V_1) < \hat{J}'(\hat{V})$. Since $J'(V_1)$ is equal to $\hat{J}'(V_1)$ and $\hat{J}$ is strictly concave, $V_1 > \hat{V}$.

Now, suppose that $V_0 < V_C$. Since $V_0 < V_C$, Proposition 1 guarantees that $J'(V_0)$ is equal to $\hat{J}'(V_C)$. Since $\hat{V} > V_C$ and $\hat{J}$ is strictly concave, it follows that $\hat{J}'(\hat{V}) < \hat{J}'(V_C)$. Putting these observations together yields $\hat{J}'(\hat{V}) < J'(V_0)$. Since $\hat{J}'(\hat{V}) < J'(V_0)$, the first necessary condition in (2.55) implies

$$R + \frac{\beta}{1-\beta}\left[J(V_1) - J(V_0)\right] > 0. \tag{2.58}$$

Using (2.58) in the second necessary condition in (2.55) yields $J'(V_1) < \hat{J}'(\hat{V})$. Since $\hat{J}'(\hat{V})$ is equal to $J'(\hat{V})$ and J is weakly concave, $V_1 > \hat{V}$. Moreover, since $V_0 < V_C$ and $\hat{V} > V_C$, we also have $V_0 < \hat{V}$.

Similar arguments can be used to prove that $V_0 < \hat{V} < V_1$ even when the agent's continuation values are not interior. ■

Proof of part (2). Since V_0 is a continuous function of $\hat{V}$ and it is such that $V_0 = V_F$ for $\hat{V} = V_F$ and $V_0 \leq \hat{V}$ for all $\hat{V} \in [V_C, V_F]$, there exists an interval (V^a, V_F) such that $V_0 \in (V_C, V_F)$ for all $\hat{V} \in (V^a, V_F)$. Hence, for all $\hat{V} \in (V^a, V_F)$, the first-order condition for V_0 is given by

$$\frac{\hat{J}'(V_0) - \hat{J}'(\hat{V})}{\hat{J}'(\hat{V})} = \frac{1}{\hat{J}'(\hat{V})}\frac{F'(R)}{F(R)}\left\{R + \frac{\beta}{1-\beta}\left[J(V_1) - J(V_0)\right]\right\}, \tag{2.59}$$

where the above expression is obtained by dividing (2.44) by $\hat{J}'(\hat{V})$ and using the fact that $J'(V_0) = \hat{J}'(V_0)$ and $\lambda = -\hat{J}'(\hat{V})$.

On the way to a contradiction, suppose that $\hat{J}'(\hat{V})$ converges to some κ for $\hat{V} \to V_F$. Since V_0 converges to V_F, $\hat{J}'(\hat{V}_0)$ converges to κ for $\hat{V} \to V_F$, it follows that the left-hand side of (2.59) converges to 0. Since R converges to η for $\hat{V} \to V_F$, it follows that the right-hand side of (2.59) converges to $(F'(\eta)\eta)/(F(\eta)\kappa) \neq 0$. We have thus reached a contradiction. Therefore, $\hat{J}'(\hat{V})$ must diverge for $\hat{V} \to V_F$. Since $\hat{J}'(\hat{V}) < 0$ for all $\hat{V} > V^*$, $\hat{J}'(\hat{V}) = -\infty$ for $\hat{V} \to V_F$.

For any $\hat{V} \in [V_C, V_F)$, $V_1 = V_F$ is optimal only if

$$\hat{J}'(V_F) - \hat{J}'(\hat{V}) \geq -\frac{F'(R)}{1-F(R)}\left\{R + \frac{\beta}{1-\beta}\left[J_F - J(V_0)\right]\right\}. \tag{2.60}$$

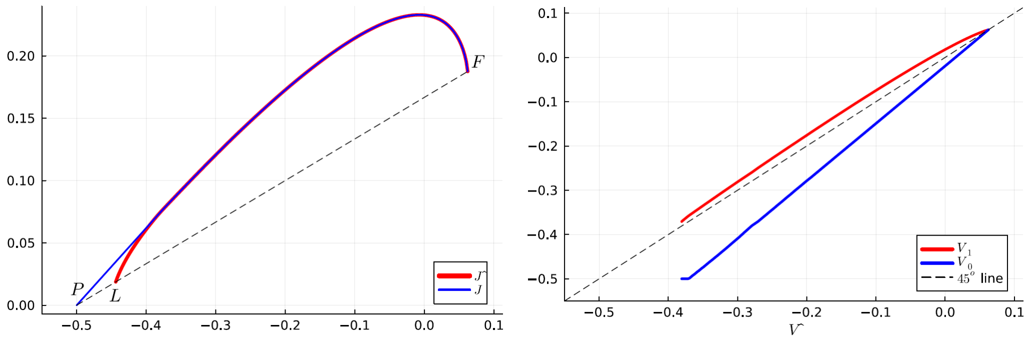
The left-hand side of (2.60) is equal to $-\infty$ since $\hat{J}'(V_F) = -\infty$ and $\hat{J}'(\hat{V})$ is finite. The right-hand side of (2.60) is finite, since $R \in (\underline{x}, \overline{x})$, $F'(R)/(1-F(R))$ is finite, and $J_F - J(V_0)$ is bounded between $J_F - J^*$ and $J_F - J_P$. Hence, the condition for the optimality of $V_1 = V_F$ cannot hold for any $\hat{V} \in [V_C, V_F)$. ■

The characterization results in Propositions 1, 2 and 3 allow us to understand the properties of the optimal mechanism. The optimal mechanism is such that initial value to the principal is J^*, with $J^* > J_F$ and $J^* > J_P$, and the initial value to the agent is V^*, with $V^* > V_C$ and $V^* < V_F$. Since $J^* > J_P$, the optimal mechanism allows the principal to achieve a value higher than what it could obtain by making the hiring decision without any input from the agent. Since $J^* > J_F$, the optimal mechanism allows the principal to achieve a value higher than what it could obtain by permanently delegating the hiring decision to the agent. The optimal mechanism provides the principal with a value higher than both J_F and J_P by using the threat of taking the hiring decision away from the agent. The threat induces the agent to follow hiring criteria that are not the agent's most preferred ones, but are tilted towards the preferences of the principal.

Since $V^* > V_C$, the mechanism starts with the hiring decision being delegated to the agent. The agent chooses to hire applicants with a quality x that is greater than some reservation R, where R is strictly smaller than η. The agent adopts a reservation quality that is tilted away from the one that maximizes its own flow payoff (η) and towards the one that maximizes the flow payoff of the principal (0) because the mechanism rewards the agent for hiring, and punishes it for not hiring. The reward to the agent is an increase in value. The punishment to the agent is a decrease in value. The increase in value moves the agent further away from the threshold V_C, below which the hiring decision may be permanently given to the principal. Thus, heuristically, the increase in value is delivered by delegating the hiring decision to the agent for a longer period of time. The increase in value also moves the agent towards a part of the mechanism where incentives are weaker, since the gap between V_1 and V_0 converges to 0 as the agent's value increases. Thus, heuristically, the increase in value is delivered by letting the agent to hire applicants according to criteria that are closer to those preferred by the agent. These properties of the optimal mechanism are illustrated in Figure 6, where we plot the expected duration of delegation and the reservation quality as a function of the agent's value. Figure 5, where we plot the value and policy functions, shows that, as conjectured, $\hat{J}(\hat{V})$ is strictly concave and that V_0 and V_1 are continuous functions of $\hat{V}$.

When the agent hires an applicant, the agent's value increases and the hiring decision is delegated to the agent for a longer period of time. Yet, no matter how many applicants the agent does hire, the hiring decision is never permanently delegated to the agent. This finding is easy to understand. If, after some history of play, the hiring decision were to be permanently delegated to the agent, the agent would hire all and only applicants with quality x greater than η and, hence, its value would be V_F. As shown in Proposition 3, however, the optimal mechanism never gives the agent a value of V_F. Since the hiring decision is never permanently delegated to the agent[5], it follows that the hiring decision is

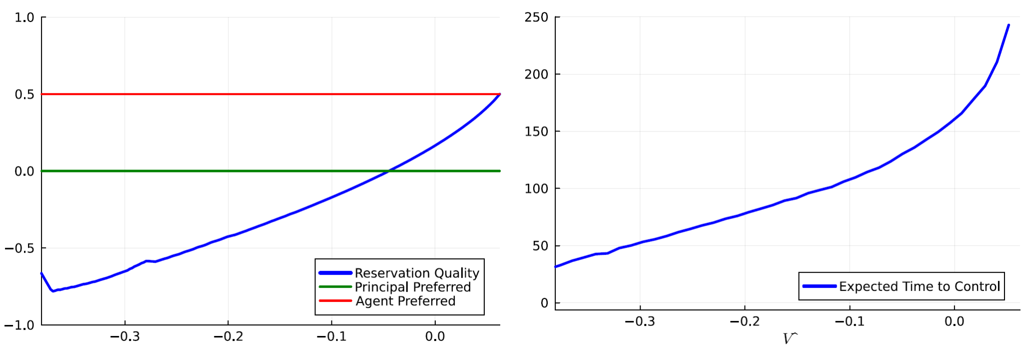
[5] The optimal mechanism never reaches V_F because $\hat{J}'(V_F)$ diverges to $-\infty$ for $V \to V_F$. In turn, this is the case because, at V_F, the agent's reservation quality R is equal to its preferred reservation quality η and different from the principal's preferred reservation quality 0. For this reason, at V_F, a small reduction



(a) Value functions (b) Policy functions

Notes: Value functions and policy functions given parameters $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on the interval $[-1, 1]$.

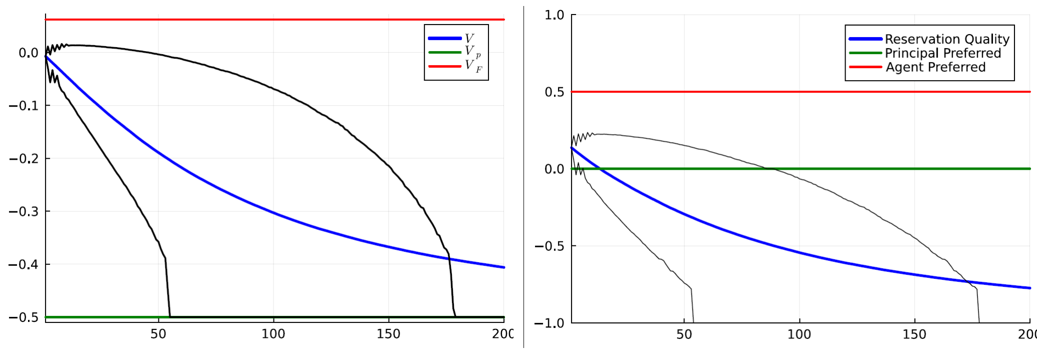
Figure 5: Numerical example: Value and policy functions



(a) Reservation quality (b) Time to control

Notes: Reservation quality R and expected time to control given parameters $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on the interval $[-1, 1]$.

Figure 6: Numerical example: Reservation quality and time to control



(a) Agent's value over time (b) Reservation quality over time

Notes: Average , $25th$ percentile and $75th$ percentile of the agent's value and the reservation quality across 100,000 simulations as a function of time, given parameters $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on the interval $[-1, 1]$.

Figure 7: Numerical example: Agent's value and reservation quality over time

eventually given to the principal. In other words, the probability that the hiring decision is given to the principal is 1 and, hence, the agent's value reaches its minimum V_P with probability 1. The finding, which is illustrated in Figure 7, is reminiscent of the "immiseration in the limit" result in the repeated moral-hazard problems studied by Thomas and Worrall (1990) and Atkeson and Lucas (1992), although the environment and the economic forces behind the result are quite different.

When the agent does not hire an applicant, the agent's value declines. Eventually, the agent's value is bound to fall below the threshold V_C. When this happens, the mechanism involves a lottery. If the lottery breaks against the agent, the hiring decision is permanently given to the principal. In this case, the value to the agent reaches its minimum V_P. If the lottery breaks in favor of the agent, the current hiring decision remains in the hands of the agent. In this case, the value to the agent moves to the threshold V_C. Then, if the agent does not hire the applicant in the current period, the hiring decision is permanently given to the principal with some positive probability. If the agent hires the applicant in the current period, the agent is rewarded by the mechanism and its value increases above the threshold V_C. The agent thus retains (temporarily) control of the hiring process.

The optimal mechanism involves ex-post inefficiencies. The most obvious of these inefficiencies occurs when the hiring decision is given to the principal. Indeed, both the agent and the principal would be better off if the hiring decision was permanently delegated to the agent rather than permanently controlled by the principal, as the value to the agent would be V_F rather than V_P and the value to the principal would be J_F rather than J_P. More generally, whenever the optimal mechanism reaches a point on the upward sloping part of the (V, J) frontier, there exists a different mechanism that is

in the agent's reservation quality R leads to a second-order decline in the agent's value and a first-order increase in the principal's value. Carrasco et al. (2017) use a similar argument.

feasible and would give both parties strictly higher values. Yet, the optimal mechanism needs these ex-post inefficiencies to deliver the highest ex-ante value to the principal. Indeed, if a mechanism never gave control of hiring to the principal, the highest ex-ante value it could deliver to the principal would be J_F. Even in the delegation phase, the optimal mechanism may feature some ex-post inefficiencies. Figure 6 shows that, when the agent's value is sufficiently low, the optimal mechanism induces the agent to use a reservation quality R that is negative. That is, when the agent's value is low, the optimal mechanism may induce the agent to hire candidates that not even the principal would want to hire.

3 Optimal control

In Section 2, we restricted attention to mechanisms such that: (i) if the mechanism hands the control of the hiring process to the principal, it does so forever; (ii) if the mechanism hands control of the hiring process to the principal, the mechanism instructs the principal to hire every applicant. In this section, we show that these seemingly arbitrary restrictions to the mechanism space are without loss in generality.

Let us start by relaxing restriction (i) and consider mechanisms that allow for the hiring decision to be either controlled by the principal or delegated to the agent in every period and after every history. We maintain the restriction to mechanisms that require the principal to hire the applicant whenever the principal controls the hiring decision. As in Section 2, the principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} \Gamma(V) = \quad & \max_{p,\tilde{V},\hat{V}} p\tilde{J}(\tilde{V}) + (1-p)\hat{J}(\hat{V}), \\ & \text{s.t. } V = p\tilde{V} + (1-p)\hat{V}, \\ & \qquad p \in [0,1], \quad \tilde{V} \in \mathcal{V}, \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{3.1}$$

where $\tilde{V}$ and $\tilde{J}(\tilde{V})$ denote the second-stage values to the agent and the principal conditional on the hiring decision being controlled by the principal in the current period, and $\hat{V}$ and $\hat{J}(\hat{V})$ denote the second-stage values to the agent and the principal conditional on the hiring decision being delegated to the agent in the current period. The value $\hat{J}(\hat{V})$ is given by (2.4). The value $\tilde{J}(\tilde{V})$ is given by

$$\begin{aligned} \tilde{J}(\tilde{V}) \quad &= (1-\beta)\int_{\underline{x}}^{\overline{x}} x dF(x) + \beta\Gamma(V_+) \\ &= (1-\beta)J_P + \beta\Gamma(V_+), \end{aligned} \tag{3.2}$$

where

$$\begin{aligned}\tilde{V} &= (1-\beta)\int_{\underline{x}}^{\overline{x}}(x-\eta)dF(x)+\beta V_{+} \\ &= (1-\beta)V_P+\beta V_{+}.\end{aligned} \tag{3.3}$$

Consider some arbitrary value V_0 of the first-stage problem to the agent. Using (3.2) to substitute $\tilde{J}(\tilde{V})$ on the right-hand side of (3.1), we can write $\Gamma(V_0)$ as

$$\Gamma(V_0)=p_0(1-\beta)J_P+(1-p_0)\hat{J}(\hat{V}_0)+p_0\beta\Gamma(V_1), \tag{3.4}$$

where p_0, $\tilde{V}_0$ and $\hat{V}_0$ are the solution to (3.1) for $V=V_0$, and V_1 is obtained from (3.3) and hence it is such that $\tilde{V}_0=(1-\beta)V_P+\beta V_1$. Note that the coefficients in front of J_P, $\hat{J}(\hat{V}_0)$ and $\Gamma(V_1)$ are positive and they sum to 1.

Using (3.1) to substitute $\Gamma(V_1)$ on the right-hand side of (3.4) and (3.2) to substitute $\tilde{J}(\tilde{V})$ in (3.1), we can write $\Gamma(V_0)$ as

$$\begin{aligned}\Gamma(V_0) &= p_0(1-\beta)J_P+\beta p_0p_1(1-\beta)J_P \\ &+(1-p_0)\hat{J}(\hat{V}_0)+\beta p_0(1-p_1)\hat{J}(\hat{V}_1)+\beta^2p_0p_1\Gamma(V_2),\end{aligned} \tag{3.5}$$

where p_1, $\tilde{V}_1$ and $\hat{V}_1$ are the solution to (3.1) for $V=V_1$, and $\tilde{V}_1$ is obtained from (3.3) and hence it is such that $\tilde{V}_1=(1-\beta)V_P+\beta V_2$. Again, the coefficients in front of J_P, $\hat{J}(\hat{V}_0)$, $\hat{J}(\hat{V}_1)$ and $\Gamma(V_2)$ are positive and they sum to 1. This is because, when we replace $\Gamma(V_1)$ on the right-hand side of (3.4) with $p_1\left[(1-\beta)J_P+\beta\Gamma(V_2)\right]+(1-p_1)\hat{J}(\hat{V}_1)$, the coefficients on J_P, $\Gamma(V_2)$ and $\hat{J}(\hat{V}_1)$ are positive and they sum to 1.

Repeating the same steps as above T times, we can write $\Gamma(V_0)$ as

$$\begin{aligned}\Gamma(V_0) &= \left[(1-\beta)\sum_{t=0}^{T}\beta^t\left(\prod_{i=0}^{t}p_i\right)\right]J_P \\ &+\sum_{t=0}^{T}\left[\beta^t(1-p_t)\left(\prod_{i=0}^{t-1}p_i\right)\hat{J}(\hat{V}_t)\right]+\left(\beta^{T+1}\prod_{i=0}^{T}p_i\right)\Gamma(V_{T+1}).\end{aligned} \tag{3.6}$$

For $T\to\infty$, the expression above becomes

$$\begin{aligned}\Gamma(V_0) &= \left[(1-\beta)\sum_{t=0}^{\infty}\beta^t\left(\prod_{i=0}^{t}p_i\right)\right]J_P \\ &+\sum_{t=0}^{\infty}\left[\beta^t(1-p_t)\left(\prod_{i=0}^{t-1}p_i\right)\hat{J}(\hat{V}_t)\right].\end{aligned} \tag{3.7}$$

Again, the coefficients in front of J_P and $\hat{J}(\hat{V}_t)$ are positive and sum up to 1.

Note that we can write (3.7) as

$$\begin{aligned}\Gamma(V_0) &= \overline{p}J_P+(1-\overline{p})\sum_{t=0}^{\infty}\left[\frac{\beta^t(1-p_t)\left(\prod_{i=0}^{t-1}p_i\right)}{1-\overline{p}}\hat{J}(\hat{V}_t)\right] \\ &\leq \overline{p}J_P+(1-\overline{p})\hat{J}\left(\overline{V}\right), \\ &\leq J(V_0),\end{aligned} \tag{3.8}$$

where $\overline{p}$ and $\overline{V}$ are given by

$$\overline{p} = (1-\beta)\sum_{t=0}^{\infty}\beta^t\left(\prod_{i=0}^{t} p_i\right) \text{ and } \overline{V} = \frac{\beta^t(1-p_t)\left(\prod_{i=0}^{t-1} p_i\right)}{1-\tilde{p}}\hat{V}_t. \tag{3.9}$$

The second line in (3.8) makes use of the fact that $\hat{J}$ is concave. The third line in (3.8) makes use of the fact that the lottery $(p, \hat{V}) = (\overline{p}, \overline{V})$ is a feasible choice in the first-stage problem (2.1) for $V = V_0$.

The inequality in the third line of (3.8), $\Gamma(V_0) \leq J(V_0)$, shows that a mechanism that allows the hiring decision to be assigned to either the agent or the principal in every period and after every history cannot give a higher value to the principal than a mechanism that, when it assigns the hiring decision to the principal, it does so forever. Since a mechanism that assigns the hiring decision to the principal once and for all is a special case of a mechanism that allows the hiring decision to be assigned to either party in every period, it follows that $\Gamma(V_0) \geq J(V_0)$. Combining these observations yields $\Gamma(V_0) = J(V_0)$. That is, the optimal mechanism that allows the hiring decision to be given to either party in every period is payoff-equivalent to the optimal mechanism in which control by the principal is permanent. The intuition for this finding is simple and can be seen from the analysis above. Namely, a lottery between temporary control by the principal and delegation to the agent is equivalent to a lottery between permanent control by the principal and delegation to the agent that attaches a lower probability on control.

Next, we relax restriction (ii) by considering mechanisms that, when the principal is in control of the hiring decision, are allowed to instruct the principal to either hire or not hire the applicants. We maintain the restriction to mechanisms that hand the hiring decision to the principal on a permanent basis. As in Section 2, the principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} \Gamma(V) = \quad & \max_{p,q,\hat{V}} pJ_P + qJ_Q + (1-p-q)\hat{J}(\hat{V}), \\ & \text{s.t. } V = pV_P + qV_Q + (1-p-q)\hat{V}, \\ & \qquad p, q \in [0,1], \quad p+q \leq 1, \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{3.10}$$

where V_Q and J_Q are the values to the agent and the principal if the principal has control over hiring and he is prescribed not hire the applicants, i.e. $(V_Q, J_Q) = (0, 0)$, V_P and J_P are the values to the agent and the principal if the principal has control over hiring and is prescribed to hire the applicants, i.e. $(V_P, J_P) = (-\eta, 0)$, and $\hat{V}$ and $\hat{J}(\hat{V})$ are the values to the agent and the principal if the hiring decision is delegated to the agent in the current period.

Let p_0, q_0 and $\hat{V}_0$ denote a solution to (3.10) for V equal to some arbitrary V_0. Suppose that $V_Q \geq V_L$. Then, the value $\Gamma(V_0)$ to the principal is such that

$$\begin{aligned}\Gamma(V_0) &= p_0 J_P + q_0 J_Q + (1 - p_0 - q_0)\hat{J}(\hat{V}_0) \\ &\leq p_0 J_P + (1 - p_0)\left[\frac{q_0}{1 - p_0}\hat{J}(V_Q) + \frac{1 - p_0 - q_0}{1 - p_0}\hat{J}(\hat{V}_0)\right] \\ &\leq p_0 J_P + (1 - p_0)\hat{J}\left(\overline{V}\right),\end{aligned} \tag{3.11}$$

where

$$\overline{V} = \frac{q_0}{1 - p_0}V_Q + \frac{1 - p_0 - q_0}{1 - p_0}\hat{V}_0. \tag{3.12}$$

The second line in (3.11) uses the fact that $\hat{J}(V_Q) > 0 = J_Q$. The third line in (3.11) makes use of the concavity of $\hat{J}$. Similarly, notice that the value V_0 to the agent is such that

$$V_0 = p_0 V_P + (1 - p_0)\overline{V}. \tag{3.13}$$

Note that the lottery $(p, \hat{V}) = (p_0, \overline{V})$ is a feasible choice in the first-stage problem (2.1). In turn, this implies that $J(V_0) \geq \Gamma(V_0)$. Since any mechanism that involves the principal hiring applicants when in control is a feasible choice for (3.11), it follows that $\Gamma(V_0) \geq J(V_0)$. Combining the two inequalities yields $\Gamma(V_0) = J(V_0)$.

Suppose that $V_Q < V_L$ and, hence, $\hat{V}_0 > V_Q$. Notice that the value $\Gamma(V_0)$ to the principal is such that

$$\begin{aligned}\Gamma(V_0) &= p_0 J_P + q_0 J_Q + (1 - p_0 - q_0)\hat{J}(\hat{V}_0) \\ &\leq \overline{p} J_P + (1 - \overline{p})\,\hat{J}(\hat{V}_0),\end{aligned} \tag{3.14}$$

where $\overline{p}$ is defined as

$$\overline{p} = p_0 + q_0\frac{\hat{V}_0 - V_Q}{\hat{V}_0 - V_P} \in \left[p_0, p_0 + q_0\right]. \tag{3.15}$$

The inequality in (3.14) makes use of the fact that $J_P = J_Q < \hat{J}(\hat{V}_0)$ and $1 - \overline{p} \geq 1 - p_0 - q$. Similarly, notice that the value V_0 to the agent is such that

$$\begin{aligned}V_0 &= p_0 V_P + q_0 V_Q + (1 - p_0 - q_0)\hat{V}_0. \\ &= \overline{p} V_P + (1 - \overline{p})\,\hat{V}_0.\end{aligned} \tag{3.16}$$

Therefore, the lottery $(p, \hat{V}) = (\overline{p}, \hat{V}_0)$ is a feasible choice in the first-stage problem (2.1). Hence, $J(V_0) \geq \Gamma(V_0)$. Since any mechanism that involves the principal hiring applicants when in control is a feasible choice for (3.11), it follows that $\Gamma(V_0) \geq J(V_0)$. Combining the two inequalities yields $\Gamma(V_0) = J(V_0)$.

Whether $V_Q \geq V_L$ or $V_Q < V_L$, $\Gamma(V_0)$ is equal to $J(V_0)$. That is, any optimal mechanism that, when the principal is in control of the hiring decision, is allowed to instruct the principal to either hire or not hire the applicants is payoff-equivalent to the optimal mechanism that, when the principal is in control of the hiring decision, instructs the principal to hire the applicants. This finding is also intuitive. If the mechanism gives control over hiring to the principal and instructs the principal to hire, the first-stage values to the agent and the principal, V and $J(V)$, are a convex combination between the point $P = (-\eta, 0)$ and the point $C = (V_C, J_C)$. Since $V_C > -\eta$ and $J_C > 0$, the convex combination generates a $J(V)$ such that $J(0) > 0$. If the mechanism is allowed to instruct the principal not to hire, the first-stage values to the agent and the principal, V and $J(V)$, can be further convexified with the point $Q = (0, 0)$. This convexification, however, is useless because Q lies below the (V, J) frontier (see Figure 2).

It is straightforward to combine the above arguments to obtain the following proposition.

Proposition 4 (Optimal control). *Consider the class of mechanisms that can assign the hiring decision to either the principal or the agent in any period and after any history, and that can prescribe that the principal hires or does not hire the applicant in any period in which the hiring decision is assigned to the principal. Let Γ denote the first-stage value function for this class of mechanisms and let J be the first-stage value function in the restricted class of mechanisms considered in Section 2. Then, $\Gamma = J$.*

4 Extensions

In this section, we use the analysis of the baseline model to characterize the optimal mechanism in richer and more realistic environments. In Section 4.1, we consider an environment in which the agent has the option to hire either a contentious applicant, an applicant from a demographic group against which the agent is biased, or a non-contentious applicant, an applicant from a demographic group against which the agent holds no bias. In Section 4.2, we consider an environment in which the agent has the option to hire one among n contentious applicants, and an environment in which the agent has the option to hire one among n contentious applicants and m uncontentious applicants. In Section 4.3, we consider a version of the baseline model in which the agent is positively biased. In Section 4.4, we consider an environment in which the principal cannot observe whether an applicant is available or not. The analysis of these extensions follows directly from the analysis of the baseline model.

4.1 Contentious and uncontentious applicants

In the baseline model of Section 2 we considered an environment in which there is only one applicant that can be hired in each period and the agent is biased against the applicant. This environment describes well a situation in which the principal operates a technology with constant returns to scale in labor and, hence, applicants do not compete against each other but only against the option of not hiring. Moreover, since applicants do not compete against each other under constant returns to scale, it is natural to restrict attention to applicants against which the agent is biased. The assumption of constant returns to scale in labor is made by most search-theoretic models of the labor market, such as Mortensen and Pissarides (1994), Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Menzio and Shi (2011). In other situations, however, it may be more realistic to assume that applicants compete against each other because the principal operates a technology with decreasing returns to scale. A simple formulation of this situation is an environment in which multiple applicants compete for a single vacancy. This is the assumption in the search-theoretic models by Montgomery (1991) and Burdett, Shi and Wright (2001). In this case, it is also natural to assume that the agent is biased against some applicants but not against others based on the applicant's demographics. In this subsection, we characterize the optimal mechanism in this alternative environment.

We consider a version of the model in which in there are multiple applicants for each vacancy. Suppose that applicants come from two different groups. Some applicants come from group X, and some applicants come from group Y. The quality x of an X-applicant is drawn from a distribution $F_x(x)$, with mean 0 and support $[\underline{x}, \overline{x}]$. The quality y of a Y-applicant is drawn from a distribution $F_y(y)$, with mean 0 and support $[\underline{y}, \overline{y}]$. The agent is negatively biased towards X-applicants, with a bias given by $\eta \in (0, \overline{x})$. The agent is unbiased towards Y-applicants. In this sense, X-applicants are contentious and Y-applicants are not contentious. The group of a particular applicant is known to both the principal and the agent, as it may reflect some readily observable demographic characteristics. The quality of a particular applicant is known only to the agent.

In order to keep the analysis simple, let us assume that there is exactly one X-applicant and one Y-applicant for each vacancy[6] and that the vacancy needs to be filled.[7] It is useful to denote as z the difference between the quality x of an X-applicant and the quality y of a Y-applicant. That is, z equals $x - y$. The joint distribution of the random variables y and z is determined by the quality distribution F_x and F_y. For our purposes, it is useful to describe the joint distribution of y and z with the marginal distribution $F_z(z)$ of the random variable z and the distribution $G(y|z)$ of the random variable y conditional on z. Note that the support of z is $[\underline{z}, \overline{z}]$, with $\underline{z} = \underline{x} - \overline{y}$ and $\overline{z} = \overline{x} - \underline{y}$.

The mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} J(V) = \quad & \max_{p, \hat{V}} pJ_P + (1-p)\hat{J}(\hat{V}) \\ & \text{s.t. } V = pV_P + (1-p)\hat{V}, \\ & \qquad p \in [0,1], \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{4.1}$$

[6] In the next subsection, we extend the analysis to the case of multiple contentious and non-contentious applicants.

[7] If the vacancy does not need to be filled, the structure of the mechanism design problem is qualitatively different than in the baseline. In particular, in the second-stage problem, the mechanism can specify different continuation values for the agent depending on whether the agent hired the contentious applicant, the agent hired the uncontentious applicant, or the agent did not hire anyone. We leave the characterization of this mechanism design problem to future work.

where V_P and J_P are respectively given by

$$V_P = -\eta, \quad J_P = 0. \tag{4.2}$$

For the same reasons as in Section 3, it is without loss in generality to restrict attention to mechanisms such that: (i) when the mechanism gives control to the principal, it does so forever; (ii) when the mechanism gives control to the principal, the mechanism instructs the principal to hire the contentious applicant. For this reason, V_P is equal to $-\eta$ and J_P is equal to 0.

At the second stage, the value of the mechanism to the agent is

$$\hat{V} = \int_z \left[\int_y \max \left\{ (1-\beta)(y+z-\eta) + \beta V_1, (1-\beta) y + \beta V_0 \right\} dG_y(y|z) \right] dF_z(z). \tag{4.3}$$

The above expression is easy to understand. Consider a particular realization of the random variables y and z. If the agent hires the X-applicant, the agent's flow payoff is $(1-\beta)(y+z-\eta)$, where $y+z$ is the quality x of the X-applicant, and the agent's continuation value is V_1. If the agent does not hire the X-applicant, the agent's flow payoff is $(1-\beta)y$, where y is the quality of the Y-applicant, and the agent's continuation value is V_0. The agent chooses whether to hire or not hire the X-applicant so as to maximize the sum of its flow and continuation payoffs. Hence, the agent hires the X-applicant if and only if $z \geq R$, where

$$R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \tag{4.4}$$

Using the definition of the reservation quality R in (4.4), we can rewrite (4.3) as

$$\begin{aligned} \hat{V} &= \int^R \left[\int_y [(1-\beta)y + \beta V_0] \, dG_y(y|z) \right] dF_z(z) \\ &+ \int_R \left[\int_y [(1-\beta)(y+z-\eta) + \beta V_1] \, dG_y(y|z) \right] dF_z(z) \\ &= (1-\beta) \int_R (z-\eta) \, dF_z(z) + \beta F_z(R) V_0 + \beta (1 - F_z(R)) V_1, \end{aligned} \tag{4.5}$$

where the last line makes use of the fact that the mean of the random variable y is 0.

At the second stage, the value of the mechanism to the principal is

$$\begin{aligned} \hat{J} &= \int^R \left[\int_y [(1-\beta)y + \beta J(V_0)] \, dG_y(y|z) \right] dF_z(z) \\ &+ \int_R \left[\int_y [(1-\beta)(y+z) + \beta J(V_1)] \, dG_y(y|z) \right] dF_z(z) \\ &= (1-\beta) \int_R z dF_z(z) + \beta F_z(R) J(V_0) + \beta (1 - F_z(R)) J(V_1). \end{aligned} \tag{4.6}$$

The above expression is also easy to understand. Consider a particular realization of the random variables y and z. If $z \geq R$, the agent hires the X-applicant. The principal's flow payoff is $(1-\beta)(y+z)$, where $y+z$ is the quality x of the X-applicant, and the principal's continuation value is $J(V_1)$. If $z < R$, the agent does not hire the X-applicant. The principal's flow payoff is $(1-\beta)y$, where y is the quality of the Y-applicant, and the principal's continuation value is $J(V_0)$.

Using (4.5) and (4.6), we can write the second-stage problem as

$$\begin{aligned} \hat{J}(\hat{V}) = & \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R z dF_z(z) + \beta \left[F_z(R) J(V_0) + (1 - F_z(R)) J(V_1) \right], \\ & \text{s.t. } \hat{V} = (1-\beta) \int_R (z - \eta)\, dF_z(z) + \beta F_z(R) V_0 + \beta \left(1 - F_z(R)\right) V_1 \qquad (4.7) \\ & \qquad R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \end{aligned}$$

Following the same argument as in Lemma 1, it is easy to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1-\beta)V_F + \beta V_P, \quad V_F \equiv \int_\eta (z-\eta) dF_z(z). \tag{4.8}$$

Following the same argument as in Lemma 2, it is east to show that

$$J^* \equiv \max_{V \in \mathcal{V}} J(V) > \int_\eta z dF_z(z) \equiv J_F \tag{4.9}$$

if the marginal distribution F_z and the bias η are such that

$$\frac{\eta F_z'(\eta)}{F_z(\eta)} > \frac{\int_\eta z dF_z(z)}{\int_\eta z dF_z(z) + \eta F(\eta)}. \tag{4.10}$$

Note that the mechanism design problem with contentious and uncontentious applicants is identical to the mechanism design problem with a single contentions applicant, except that the distribution F_z of the difference in quality between the contentious and the uncontentious applicants takes the place of the distribution F of the quality of the contentious applicant. Since the distribution F_z has a mean of 0, just like the distribution F, the mechanism design problem with contentious and uncontentious applicants is a special case of the mechanism design problem with a single contentious applicant. Therefore, under condition (4.10), we can directly apply Propositions 1, 2 and 3.

Proposition 5: (Contentious and uncontentious applicants) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.1) is a lottery $(p, \hat{V})$ such that $p > 0$ and $\hat{V} = V_C$ for all $V \in [V_P, V_C)$, and such that $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, where V_C is given by (2.34).*
2. *The solution to the second-stage problem in (4.5) is a pair of agent's continuation*

values (V_0, V_1) such that $V_0 \leq V_C$, $V_1 \in [V_C, V_F)$ and $V_0 < V_1$ for $\hat{V} = V_C$, such that $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$, and such that $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$. For all $\hat{V} \in [V_C, V_F)$, the solution induces the agent to follow a threshold R on the difference between the quality of the contentious and the uncontentious applicants such that $R \in (\underline{z}, \eta)$.

The version of the model in which a contentious and an uncontentious applicant are available to fill each vacancy and each vacancy needs to be filled is isomorphic to the baseline model, except that the distribution F needs to be interpreted as the distribution of the quality gap between the contentious and the uncontentious applicants and not hiring needs to be interpreted as hiring the uncontentious applicant. It is useful to read Figure 6 in light of this alternative interpretation. The figure shows that the reservation quality gap is smaller than η, it is strictly positive when the agent's promised value is sufficiently high, and it is strictly negative when the agent's promised value is sufficiently low. This implies that, when the agent's promised value is sufficiently high, the agent's hiring decision is still biased against the contentious applicant and in favor of the uncontentious applicant, although the extent of the bias is less than under unfettered delegation. When the agent's promised value is sufficiently low, the agent's hiring decision becomes biased in favor of the contentious applicant and against the uncontentious applicant, since the agent's reservation quality gap becomes negative. Therefore, the optimal mechanism is such that, after some histories, the agent is induced to switch the direction of its bias. Even though this is ex-post inefficient, it is ex-ante optimal.

4.2 Multiple applicants

In this subsection, we characterize the optimal mechanism in a version of the baseline model in which multiple contentious applicants and multiple uncontentious applicants compete for each of vacancy. It is useful to start the analysis from the case in which there are $n \in \mathbb{N}$ contentious applicants and no uncontentious applicants. Each contentious applicant has a quality x that is independently drawn from the distribution $F(x)$. Let x_n denote the maximum of the quality of the n contentious applicants. Let $F_n(x_n)$ denote the distribution of x_n, i.e., $F_n(x) = F(x)^n$.

The principal's optimal mechanism design problem can be written recursively as a two-stage problem. The first-stage of the recursive problem is

$$\begin{aligned} J(V) = \quad & \max_{p, \hat{V}} pJ_P + (1-p)\hat{J}(\hat{V}) \\ & \text{s.t. } V = pV_P + (1-p)\hat{V}, \\ & \quad p \in [0,1], \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{4.11}$$

where the values V_P and J_P are given by

$$V_P = \int_{\underline{x}}^{\overline{x}} (x_n - \eta) dF_n(x_n), \quad J_P = \int_{\underline{x}}^{\overline{x}} x_n dF_n(x_n). \tag{4.12}$$

The second-stage of the recursive problem is

$$\begin{aligned} \hat{J}(\hat{V}) = \quad & \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R x_n dF_n(x_n) + \beta \left[F_n(R) J(V_0) + (1 - F_n(R)) J(V_1) \right], \\ & \text{s.t. } \hat{V} = (1-\beta) \int_R (x_n - \eta) dF_n(x) + \beta \left[F_n(R) V_0 + (1 - F_n(R))) V_1 \right] \\ & \quad R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \end{aligned} \tag{4.13}$$

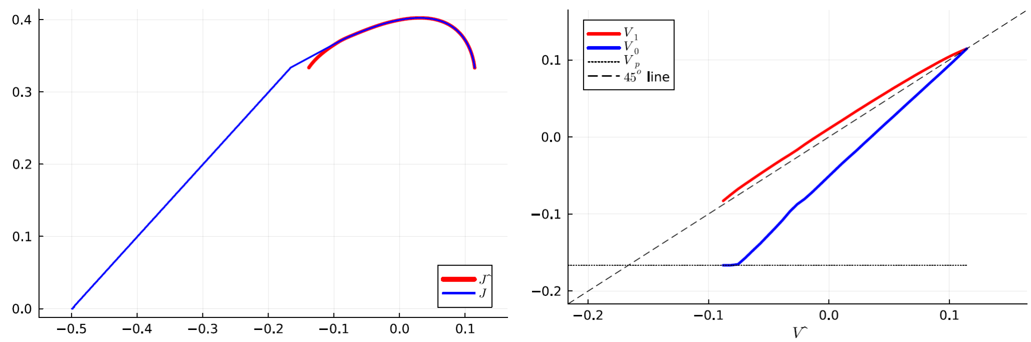
The formulation of (4.11) subsumes a restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism instructs the agent to report the ranking of the applicants, which the agent has an incentive to disclose truthfully, and then it instructs the principal to hire the highest-ranked applicant. The first restriction on the mechanisms is without loss in generality. The second restriction may not be without loss in generality. If the principal has control over hiring and hires a randomly-selected applicant, the values to the agent and the principal are $-\eta$ and 0. If the principal has control over hiring and hires the best applicant, the value to the agent is $V_P = E[x_n] - \eta$ and the value to the principal is $J_P = E[x_n]$. The optimal lottery between delegation and control may be a convex combination between a delegation point (V_C, J_C) and the control point $(-\eta, 0)$. Alternatively, the optimal lottery may be a convex combination between a delegation point (V_C, J_C) and the control point (V_P, J_P) for $V \in [V_P, V_C)$ and a convex combination between the control points $(-\eta, 0)$ and (V_P, J_P) for $V \in [-\eta, V_P)$. If the optimal mechanism in delegation is such that $V_0 \geq V_P$ for all $\hat{V} \geq V_C$, the only relevant control is the point (V_P, J_P) and the restriction (ii) is innocuous.[8] This is the case in the numerical example illustrated in Figure 8 and in all the other numerical examples we explored.

Taking as given the restrictions on the space of mechanisms implicit in (4.11), it is immediate to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1-\beta) V_F + \beta V_P, \quad V_F \equiv \int_\eta (x_n - \eta) dF_n(x_n). \tag{4.14}$$

Similarly, it is straightforward to show that J^*, which is defined as the maximum of $J(V)$ with respect to V, is strictly greater than J_F, which is defined as $\int_\eta x_n dF_n(x_n)$, if

[8] The analysis of the optimal mechanism when the optimal control involves hiring an applicant at random is similar, and is omitted for the sake of brevity.



(a) Value functions (b) Policy functions

Notes: Value functions and policy functions given parameters $n = 2$, $\beta = 0.9$, η= 0.5, and $F(x)$ uniform on the interval $[-1, 1]$.

Figure 8: Numerical example: Value and policy functions with n applicants

the quality distribution F_n and the bias η are such that

$$\frac{\eta F_n'(\eta)}{F_n(\eta)} > \frac{\int_\eta x_n dF_n(x_n) - \int_{\underline{x}}^{\overline{x}} x_n dF_n(x_n)}{\int_\eta (x_n - \eta) dF_n(x_n) - \int_{\underline{x}}^{\overline{x}} (x_n - \eta) dF_n(x_n)}. \tag{4.15}$$

Note that the mechanism design with multiple contentious applicants is identical to the mechanism design problem with a single applicant, except that the distribution F_n of the highest quality x_n among n applicants replaces the distribution F of the quality of a single applicant. The only substantive difference between the distribution F_n and the distribution F is that the former has a strictly positive mean and the latter has a mean equal to 0. Yet, under condition (4.15), it is easy to verify that the characterization results in Propositions 1, 2 and 3 also apply to the case of a distribution with a strictly positive mean.

Proposition 6: (Multiple contentious applicants) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.11) is a lottery $(p, \hat{V})$ such that $p > 0$ and $\hat{V} = V_C$ for all $V \in [V_P, V_C)$, and such that $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, where V_C is given by (2.34).*
2. *The solution to the second-stage problem in (4.13) is a pair of continuation values (V_0, V_1) such that $V_0 \leq V_C$, $V_1 \in [V_C, V_F)$ and $V_0 < V_1$ for $\hat{V} = V_C$, such that $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$, and such that $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$. For $\hat{V} \in [V_C, V_F)$, the solution induces an agent's reservation quality R such that $R \in (\underline{x}, \eta)$.*

In light of the results of the previous subsection, it is easy to extend Proposition 6 to the case in which $n \in \mathbb{N}$ contentious applicants and $m \in \mathbb{N}$ non-contentious applicants compete for each vacancy, and each vacancy needs to be filled. Let x_n denote the maximum quality of the contentious applicants, and let $F_{x,n}(x_n)$ denote the distribution of x_n. Let y_m denote the maximum quality of uncontentious applicants, and let $F_{y,m}(y_m)$ denote the distribution of y_m. Let z denote the difference between the maximum quality x_n of the contentious applicants and the maximum quality y_m of the uncontentious applicants, and let $F_z(z)$ denote the distribution of the random variable z.

The mechanism design problem can be written recursively as a two-stage problem.[9] The first-stage problem is (4.11), where V_P and J_P are respectively given by

$$V_P = \mu + \int_{\underline{z}}^{\overline{z}} (z-\eta) dF_z(z), \quad J_P = \mu + \int_{\underline{z}}^{\overline{z}} z dF_z(z). \tag{4.16}$$

and μ denotes the unconditional mean of the random variable y_m and is given by

$$\mu = \int y_m dF_{y,m}(y_m). \tag{4.17}$$

The second-stage problem is

$$\begin{aligned} \hat{J}(\hat{V}) = \quad & \max_{V_0, V_1 \in \mathcal{V}} (1-\beta)\left[\mu + \textstyle\int_R z dF_z(z)\right] + \beta\left[F_z(R)J(V_0) + (1-F_z(R))J(V_1)\right], \\ & \text{s.t. } \hat{V} = (1-\beta)\left[\mu + \textstyle\int_R (z-\eta)\, dF_z(z)\right] + \beta F_z(R)V_0 + \beta\left(1-F_z(R)\right)V_1 \\ & R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \end{aligned} \tag{4.18}$$

The implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1-\beta)V_F + \beta V_P, \quad V_F \equiv \mu + \int_\eta (z-\eta) dF_z(z). \tag{4.19}$$

The optimal mechanism is not unfettered delegation, in the sense that

$$J^* \equiv \max_{V \in \mathcal{V}} J(V) > \mu + \int_\eta z dF_z(z) \equiv J_F \tag{4.20}$$

if the marginal distribution F_z and the bias η are such that

$$\frac{\eta F_z'(\eta)}{1-F_z(\eta)} > \frac{\int_\eta z dF_z(z) - \int_{\underline{z}}^{\overline{z}} z dF_z(z)}{\int_\eta (z-\eta) dF_z(z) - \int_{\underline{z}}^{\overline{z}} (z-\eta) dF_z(z)}. \tag{4.21}$$

The mechanism design with multiple contentious and multiple uncontentious appli-

[9] We maintain the restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism instructs the agent to report the ranking of the contentious applicants, and it instructs the principal to hire the highest-ranked contentious applicant.

cants is identical to the mechanism design problem with a multiple contentious applicant, except that the distribution F_z of the gap between the highest quality x_n among the n contentious applicants and the highest quality y_m among the m uncontentious applicants replaces F_n, and the payoffs to the agent and the principal are shifted by a constant μ. Propositions 1, 2 and 3 generalize to the case in which the distribution that has a non-zero mean and to the case in which the payoffs are shifted by a constant. Therefore, under condition (4.21), Propositions 1, 2 and 3 and, in turn, Proposition 6 apply to the version of the model with multiple contentious and uncontentious applicants.

4.3 Positive bias

In the baseline model of Section 2, we characterize the optimal mechanism when the agent is biased against the applicants. In this subsection, we want to characterize the optimal mechanism in an environment where the agent is biased in favor of the applicants. In particular, we consider a version of the baseline model in which the difference between the payoff to the agent and the payoff to the principal if an applicant is hired is a strictly positive constant ϕ.

The principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage of the recursive problem is

$$\begin{aligned} J_+(V) = \quad & \max_{p_+, \hat{V}_+} p_+ J_P^+ + (1-p_+)\hat{J}_+(\hat{V}_+) \\ & \text{s.t. } V = p_+ V_P^+ + (1-p_+)\hat{V}_+, \\ & \qquad p_+ \in [0,1], \quad \hat{V}_+ \in \hat{\mathcal{V}}_+, \end{aligned} \tag{4.22}$$

where the values V_P^+ and J_P^+ are given by

$$V_P^+ = 0, \quad J_P^+ = 0. \tag{4.23}$$

The second-stage of the recursive problem is

$$\begin{aligned} \hat{J}_+(\hat{V}) = \quad & \max_{V_0^+, V_1^+ \in \mathcal{V}_+} (1-\beta)\int_{R_+} x dF(x) + \beta\left[F(R_+)J(V_0^+) + (1-F(R_+))J(V_1^+)\right], \\ & \text{s.t. } \hat{V} = (1-\beta)\int_{R_+} (x+\phi)\, dF(x) + \beta[F(R_+)V_0^+ + (1-F(R_+))\, V_1^+], \\ & \qquad R_+ = -\phi - \frac{\beta}{1-\beta}(V_1^+ - V_0^+). \end{aligned} \tag{4.24}$$

The formulation of (4.22) assumes a restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism prescribes that the principal hires none of the applicants. The first restriction is without loss in generality, for the same reasons as in Section 3. The second restriction is also without loss in generality. This is intuitive. If the principal has control over hiring and hires all the applicants, the values to the agent and the principal are ϕ and 0. If the principal has control over hiring and does not hire any applicants, the value to both the agent and the principal is 0. Clearly, the optimal lottery between delegation and control is a convex combination between a point on the delegation frontier $(\hat{V}, \hat{J}_+)$ and the control point $(0, 0)$.

As in Lemma 1, we can prove that the implementable set $\mathcal{V}_+$ is given by the interval $[V_P^+, V_F^+]$, and the implementable set $\hat{\mathcal{V}}_+$ is given by the interval $[V_L^+, V_F^+]$, where

$$V_L^+ \equiv (1-\beta)V_F^+ + \beta V_P^+, \quad V_F^+ \equiv \int_{-\phi} (x+\phi)dF(x). \tag{4.25}$$

As in Lemma 2, we can prove that J_+^*, which is defined as the maximum of $J_+(V)$ with respect to V, is strictly greater than J_F^+, which is defined as $\int_{-\phi} xdF(x)$, if the quality distribution F and the bias ϕ are such that

$$\frac{\phi F'(-\phi)}{1-F(-\phi)} > \frac{\int_{-\phi} xdF(x)}{\int_{-\phi}(x+\phi)dF(x)}. \tag{4.26}$$

Under condition (4.26), we can use the same arguments as in Propositions 1, 2 and 3 to show that the optimal mechanism has the same qualitative features as in the baseline model. There are only two substantive differences between the optimal mechanism when the agent is positively biased and the optimal mechanism when the agent is negatively biased. First, as we have already mentioned, if the agent is positively biased, the optimal mechanism is such that, when the hiring decision is controlled by the principal, the principal is required to hire none of the applicants. If the agent is negatively biased, the optimal mechanism is such that, when the hiring decision is controlled by the principal, the principal is required to hire all of the applicants. Second, if the agent is positively biased, the optimal mechanism is such that, when the hiring decision is delegated to the agent, the agent is punished for hiring an applicant with a lower continuation values, and it is rewarded for not hiring an applicant with a higher continuation value. If the agent is negatively biased, the optimal mechanism is such that the agent is rewarded for hiring an applicant and punished for not hiring an applicant. This difference is easy to understand. If the agent is positively biased, the agent wants to hire more applicants than the principal. The optimal mechanism reduces the gap between the preferences of the agent and the principal by rewarding the agent for not hiring applicants and by punishing him for hiring applicants.

The proposition below contains the characterization of the optimal mechanism when the agent is positively biased.

Proposition 7: (Positive bias) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.22) is a lottery* $(p_+, \hat{V}_+)$ *such that*

$p_+ > 0$ *and* $\hat{V}_+ = V_C$ *for all* $V \in [V_P^+, V_C^+)$, *and such that* $p_+ = 0$ *and* $\hat{V}_+ = V$ *for all* $V \in [V_C^+, V_F^+]$, *where* V_C^+ *is given by (2.34).*

2. *The solution to the second-stage problem in (4.24) is a pair of continuation values* (V_0^+, V_1^+) *such that* $V_1^+ \leq V_C^+$, $V_0^+ \in [V_C^+, V_F^+)$ *and* $V_1^+ < V_0^+$ *for* $\hat{V} = V_C^+$, *such that* $V_0^+ \in (\hat{V}, V_F^+)$ *and* $V_1^+ < \hat{V}$ *for all* $\hat{V} \in (V_C^+, V_F^+)$, *and such that* $V_0^+ = V_F^+$ *and* $V_1^+ = V_F^+$ *for* $\hat{V} = V_F^+$. *For all* $\hat{V} \in [V_C^+, V_F^+)$, *the solution induces an agent's reservation quality* R *such that* $R \in (-\phi, \overline{x})$.

Under some conditions on ϕ and F, the solution to the mechanism design problem under positive bias can be expressed as a simple transformation of the solution to the mechanism design problem under negative bias. Specifically, if the quality distribution F is symmetric around 0, in the sense that $F(x) = 1 - F(-x)$ for all $x \in [\underline{x}, \overline{x}]$, and the positive bias has the same magnitude as the negative bias, in the sense that $\phi = \eta$, the value functions that solve (4.22) and (4.26) are an horizontal translation of the value functions that solve (2.1) and (2.4), and the policy functions associated with (4.22) and (4.24) are an horizontal translation of the policy functions that solve (2.1) and (2.4), with V_0^+ taking the place of V_1 and V_1^+ taking the place of V_0.

The next proposition makes the statements above formal. The proof is in the Appendix

Proposition 8: (Positive and negative bias under symmetry). *Assume* $F(x) = 1 - F(-x)$ *for all* $x \in [\underline{x}, \overline{x}]$ *and* $\phi = \eta$. *Then:*

1. *The first-stage and second-stage value functions* J_+ *and* $\hat{J}_+$ *are such that*

$$J_+(V + \eta) = J(V), \quad \hat{J}_+(\hat{V} + \eta) = \hat{J}(\hat{V}). \tag{4.27}$$

2. *The solution to the first-stage problem in (4.22) given the promised value* $V_+ = V + \eta$ *is the lottery* $(p_+, \hat{V}_+)$, *where* $p_+ = p$, $\hat{V}_+ = \hat{V} + \eta$, *and* $(p, \hat{V})$ *is the solution to the first-stage problem in (2.1) given the promised value* V.
3. *The solution to the second-stage problem in (4.24) given the promised value* $\hat{V}_+ = \hat{V} + \eta$ *is a pair of continuation values* (V_0^+, V_1^+), *where* $V_0^+ = V_1 + \eta$, *and* $V_1^+ = V_0 + \eta$, *and* (V_0, V_1) *is the solution to the second-stage problem in (2.4) given the promised value* $\hat{V}$.
4. *The agent's reservation quality associated with the solution to (4.24) given the promised value* $\hat{V}_+ = \hat{V} + \eta$ *is* $R_+ = -R$, *where* R *is the agent's reservation quality associated with the solution to (2.4) given the promised value* $\hat{V}$.

4.4 Privately observed arrival of applicants

In the baseline model, we assume that the principal knows when there is an applicant that could be hired. In some situations, though, it may be more realistic to assume that the principal does not know whether an applicant is available or not. In this subsection, we characterize the optimal mechanism when the agent privately observes whether an applicant is available or not and, in the former case, the agent privately observes the quality of the applicant. Specifically, suppose that an applicant is available with probability $\rho \in (0,1)$ and no applicant is available with probability $1-\rho$. As in the baseline model, we assume that the agent is biased against the applicant.

As usual, the mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} J(V) = \quad & \max_{p,\hat{V}} pJ_P + (1-p)\hat{J}(\hat{V}) \\ & \text{s.t. } V = pV_P + (1-p)\hat{V}, \\ & \qquad p \in [0,1], \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{4.28}$$

where V_p and J_P are respectively given by

$$V_P = 0, \quad J_P = 0. \tag{4.29}$$

The above formulation of the first-stage problem assumes that the optimal mechanism is such that: (i) if the mechanism gives control of hiring to the principal, it does so forever; (ii) when the principal has control over hiring, the mechanism instructs the principal not hire. These assumptions are without loss in generality. The first assumption is without loss in generality for the same reasons discussed in Section 3. The second assumption is without loss in generality because not hiring is the only feasible mechanism when the principal has control over hiring. In fact, as the principal does not know whether an applicant is available or not, it needs to rely on a report by the agent. Since the agent prefers not hiring an applicant than hiring an applicant irrespective of quality, the only mechanism that induces the agent to truthfully report whether an applicant is available or not is such that, irrespective of the agent's report, the principal does not hire.

The second-stage problem is

$$\begin{aligned} \hat{J}(\hat{V}) \quad &= \max_{V_0,V_1 \in \mathcal{V}} (1-\beta)\rho \int_R x dF(x) + \beta \left[(1-\rho(1-F(R)))\, J(V_0) + \rho(1-F(R)) J(V_1) \right], \\ & \text{s.t. } \hat{V} = (1-\beta)\rho \int_R (x-\eta) dF(x) + \beta \left[(1-\rho(1-F(R)) V_0 + \rho(1-F(R))) V_1 \right] \\ & \qquad R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \end{aligned} \tag{4.30}$$

The problem above is easy to understand. In the current period, an applicant is available with probability ρ. If the applicant's quality x is higher than R, the agent hires the applicant. In this case, the principal's flow payoff is x, the agent's flow payoff is $x-\eta$, the principal's continuation value is $J(V_1)$, and the agent's continuation value is V_1. If an applicant is not available or if an applicant is available and their quality x is lower than

R, the agent does not hire. In this case, the principal's and the agent's flow payoffs are 0, the principal's continuation value is $J(V_0)$, and the agent's continuation value is V_0. Clearly, since the principal does not observe whether an applicant is or is not available, the agent's continuation payoff can only be contingent on whether an applicant is hired or not.

Following the same argument as in Lemma 1, it is easy to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1-\beta)V_F + \beta V_P, \quad V_F \equiv \rho \int_\eta (x-\eta) dF(x). \tag{4.31}$$

Following the same arguments as in Lemma 2, it is easy to show that J^*, which is defined as the maximum of $J(V)$ with respect to V, is strictly greater than J_F, which is given by $\rho \int_\eta x dF(x)$, if the quality distribution F, the bias η, anf the arrival probability ρ are such that

$$\frac{\rho \eta F'(\eta)}{1-\rho(1-F(\eta))} > \frac{\rho \int_\eta x dF(x)}{\rho \int_\eta (x-\eta) dF(x)}. \tag{4.32}$$

The mechanism design problem where the arrival of applicants is privately observed by the agent is very similar to the baseline mechanism design problem in Section 2. There are three minor differences. First, the value to the agent and the principal in control are both equal to 0, rather than $-\eta$ and 0. Second, the agent and principal's flow payoffs in delegation are scaled by the factor ρ. Lastly, the probability that the agent hires an applicant in delegation is $\rho(1-F(R))$, rather than $1-F(R)$. Notwithstanding these differences, it is easy to generalize Propositions 1 and 3. Formally, under condition (4.32), we can establish the following result.

Proposition 9: (Privately observed arrival of applicants). *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.28) is a lottery $(p, \hat{V})$ such that $p > 0$ and $\hat{V} = V_C$ for all $V \in [V_P, V_C)$, and such that $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, where V_C is given by (2.34).*
2. *The solution to the second-stage problem in (4.30) is a pair of continuation values (V_0, V_1) such that $V_0 \leq V_C$, $V_1 \in [V_C, V_F)$ and $V_0 < V_1$ for $\hat{V} = V_C$, such that $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$, and such that $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$.*

5 Conclusions

In this paper, we approached the design of anti-discriminatory labor market regulations as a delegation problem. A private firm (the agent) is repeatedly faced with the opportunity of hiring one among several applicants to fill its vacancies. The firm is biased against applicants from some demographic groups, and it is neutral towards applicants from some other demographic groups. Applicants differ not only with respect to their demographic characteristics, but also with respect to the idiosyncratic quality of their match with the firm. A benevolent and unbiased labor market authority (the principal) enacts a hiring regulation (a direct-revelation mechanism) in order to reduce the impact of the firm's bias on its hiring behavior. The hiring regulation is constrained by the fact that the quality of the match between any particular applicant and the firm is privately observed by the firm.

We showed that the optimal regulation is dynamic. Depending on the history of the firm, the regulation prescribes that either the firm is free to choose whom to hire (i.e., the hiring decision is delegated to the firm), or the labor market authority chooses whom to hire (i.e., the hiring decision is controlled by the authority). In the delegation state, the regulation rewards the firm for hiring applicants from the discriminated groups and it punishes the firm for hiring applicants from the non-discriminated groups. The reward is delivered as an increase in the expected time until the hiring decision is given to the labor market authority, and as a weakening of the incentives given to the firm to hire discriminated applicants. The punishment is delivered as a reduction in the expected time until the hiring decision is taken away from the firm, and as a strengthening of the incentives given to the firm to hire discriminated applicants. In the control state, the authority hires the best applicant among those coming from discriminated groups. The control state is absorbing and it is reached with certainty.

Much more work lies head. From the perspective of delegation theory, it would be important to extend the analysis to the case in which the principal cannot directly observe the extent of the agent's bias. From the perspective of labor market discrimination, many extensions are worthwhile. In this paper, we assumed that the firm's only choice is whether or not to hire applicants. In reality, a firm also chooses whether or not to fire its employees and, for this reason, the firm may be able to circumvent the hiring regulation. Therefore, it would be important to extend the analysis to an environment in which biased firms hire and fire workers. In this paper, we assumed away all general equilibrium considerations and, in particular, we kept both the wage and the arrival rate of applicants as exogenous. It would be important to embed the mechanism design problem studied in this paper in a general equilibrium model.

References

[1] Atkeson, A., and R. Lucas. "On Efficient Distribution with Private Information." *Review of Economic Studies*, 59: 427-453.

[2] Aghion, P., and J. Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy* 105: 1-29.

[3] Alonso, R., and N. Matouschek. 2008. "Optimal Delegation." *Review of Economic Studies*, 75: 259-293.

[4] Amador, M., and K. Bagwell. 2013. "The Theory of Optimal Delegation with an Application to Tariff Caps." *Econometrica*, 81: 1541-1599.

[5] Angeletos, M., I. Werning, and M. Amador. 2006. "Commitment vs. Flexibility." *Econometrica*, 74: 365-396.

[6] Athey, S., A. Atkeson, and P. Kehoe. 2005. "The Optimal Degree of Discretion in Monetary Policy." *Econometrica*, 73: 1431-1475.

[7] Burdett, K., and D. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39: 955-970.

[8] Burdett, K., S. Shi, and R. Wright. 2001. Pricing and Matching with Frictions. *Journal of Political Economy* 109: 1060–85.

[9] Carrasco, V., W. Fuchs, and S. Fukuda. 2019. "From Equals to Despots: The Dynamics of Repeated Decision Making in Partnerships with Private Information.“ *Journal of Economic Theory*, 182: 402-432.

[10] Clementi, G., and H. Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics." *Quarterly Journal of Economics*, 121: 229-265.

[11] Escobar, J., and J. Toikka. 2013. "Efficiency in Games with Markovian Private Information." *Econometrica*, 81: 1887-1934.

[12] Frankel, A. 2016a. "Discounted Quotas." *Journal of Economic Theory,* 166: 396-444.

[13] Frankel, A. 2016b. "Delegating Multiple Decisions." *American Economic Journal: Microeconomics*, 8: 16-53.

[14] Frankel, A. 2021. "Selecting Applicants." *Econometrica*, 89: 615-645.

[15] Guo, Y., and J. Horner. 2020. "Dynamic Allocation without Money." Manuscript, Yale University.

[16] Halac, M., and P. Yared. 2018. "Fiscal Rules and Discretion in a World Economy." *American Economic Review*, 108: 2305-2334.

[17] Holmstrom, B. 1977. "On Incentives and Control in Organizations." Manuscript, Stanford University.

[18] Jackson, M., and H. Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." *Econometrica*, 75: 241-257.

[19] Li, J., N. Matouschek, and M. Powell. 2017. "Power Dynamics in Organizations." *American Economic Journal: Microeconomics*, 9: 217-241.

[20] Lipnowski, E., and J. Ramos. 2021. "Repeated Delegation." *Journal of Economic Theory*, 188: 1-41.

[21] Malenko, A. 2019. "Optimal Dynamic Capital Budgeting." *Review of Economic Studies*, 86: 1747-1778.

[22] Menzio, G., and S. Shi. 2011. "Efficient Search on the Job and the Business Cycle." *Journal of Political Economy*, 119: 468-510.

[23] Montgomery, J. 1991. Equilibrium Wage Dispersion and Interindustry Wage Differentials. *Quarterly Journal of Economics* 106: 163–79.

[24] Mortensen, D. and C. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*, 61: 397-415.

[25] Padró i Miquel, G., and P. Yared. "The Political Economy of Indirect Control." *Quarterly Journal of Economics*, 127: 947-1015.

[26] Postel Vinay, F., and J. Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70: 2295-2350.

[27] Sannikov, Y. 2008. "A Continuous-Time Version of the Principal-Agent Problem." *Review of Economic Studies*, 75: 957-984.

[28] Thomas, J., and T. Worrall. 1990. "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." *Journal of Economic Theory*, 51: 367-390.

Appendix

A Proof of Proposition 8

For $\hat{V}_+ = \hat{V} + \eta$, the functional equation (4.24) can be written

$$\begin{aligned} & \hat{J}_+(\hat{V}+\eta) \\ = \quad & \max_{V_0^+, V_1^+} (1-\beta)\int_{R_+} xdF(x) + \beta\left[F(R_+)J_+(V_0^+) + (1-F(R_+))J_+(V_1^+)\right], \\ \text{s.t.} \quad & \hat{V}+\eta = (1-\beta)\int_{R_+}(x+\phi)dF(x) + \beta\left[F(R_+)V_0^+ + (1-F(R_+)))V_1^+\right] \\ & R_+ = -\phi - \frac{\beta}{1-\beta}(V_1^+ - V_0^+), \quad V_0^+,\, V_1^+ \in [V_P^+, V_F^+]. \end{aligned} \tag{A.1}$$

Let us define V_0 and V_1 as, respectively, $V_1^+ - \eta$ and $V_0^+ - \eta$. Since the choice variables V_0^+ and V_1^+ must belong to the interval $[V_P^+, V_F^+]$, the alternative choice variables V_0 and V_1 must belong to the interval

$$\begin{aligned} \left[V_P^+ - \eta, V_F^+ - \eta\right] &= \left[-\eta, \int_{-\phi}(x+\phi)dF(x) - \eta\right] \\ &= \left[-\eta, \int_{-\phi} xdF(x) + (1-F(-\phi))\phi - \eta\right] \\ &= \left[-\eta, \int_{\eta} xdF(x) + F(\eta)\eta - \eta\right] \\ &= \left[-\eta, \int_{\eta}(x-\eta)dF(x)\right] = [V_P, V_F] \end{aligned} \tag{A.2}$$

The first line makes use of the definitions of V_P^+ and V_F^+. The third line makes use of the fact that $1 - F(-\phi)$ equals $F(\phi)$ and $\phi = \eta$, as well as of the fact that, by symmetry of the quality distribution, $\int_{-\phi} xdF(x)$ equals $\int_{\phi} xdF(x)$. The fourth line makes use of the definition of V_P and V_F.

Using the above definitions and observations, we can rewrite (A.1) as

$$\begin{aligned} & \hat{J}_+(\hat{V}+\eta) \\ = \quad & \max_{V_0, V_1}(1-\beta)\int_{R_+} xdF(x) + \beta\left[F(R_+)J_+(V_1+\eta) + (1-F(R_+))J_+(V_0+\eta)\right], \\ \text{s.t.} \quad & \hat{V}+\eta = (1-\beta)\int_{R_+}(x+\phi)dF(x) + \beta\left[F(R_+)V_1 + (1-F(R_+))V_0 + \eta\right] \\ & R_+ = -\eta - \frac{\beta}{1-\beta}(V_0 - V_1), \quad V_0,\, V_1 \in [V_P, V_F]. \end{aligned} \tag{A.3}$$

Defining R as $-R_+$ and using the fact that $F(R_+) = 1 - F(-R_+)$, we can rewrite

(A.3) as

$$
\begin{aligned}
& \hat{J}_+(\hat{V}+\eta) \\
= \quad & \max_{V_0,V_1}(1-\beta)\int_{R_+} x dF(x) + \beta\left[(1-F(R))J_+(V_1+\eta)+F(R)J_+(V_0+\eta)\right], \\
\text{s.t.} \quad & \hat{V}+\eta = (1-\beta)\int_{R_+}(x+\eta)dF(x)+\beta\left[(1-F(R))V_1+F(R)V_0+\eta)\right] \\
& R = \eta - \frac{\beta}{1-\beta}(V_1-V_0), \quad V_0,\, V_1 \in [V_P, V_F].
\end{aligned} \tag{A.4}
$$

Using the symmetry of the quality distribution and the definition of R, we can now rewrite (A.4) as

$$
\begin{aligned}
& \hat{J}_+(\hat{V}+\eta) \\
= \quad & \max_{V_0,V_1}(1-\beta)\int_{R} x dF(x) + \beta\left[(1-F(R))J_+(V_1+\eta)+F(R)J_+(V_0+\eta)\right], \\
\text{s.t.} \quad & \hat{V} = (1-\beta)\int_{R}(x-\eta)dF(x)+\beta\left[(1-F(R))V_1+F(R)V_0\right] \\
& R = \eta - \frac{\beta}{1-\beta}(V_1-V_0), \quad V_0,\, V_1 \in [V_P, V_F].
\end{aligned} \tag{A.5}
$$

In the objective function, we used the fact that $\int_{R_+} x dF(x)$ equals $\int_{-R_+} x dF(x)$ and the definition $R=-R_+$. In the promise-keeping constraint, we collected η on the right-hand side and used the fact that $\int_{R_+} x dF(x)$ equals $\int_R x dF(x)$.

For $\hat{V}+\eta$, the functional equation (4.22) can be written as

$$
\begin{aligned}
J_+(V+\eta) \quad &= \max_{p_+,\hat{V}_+} p_+ J_P^+ + (1-p_+)\hat{J}_+(\hat{V}_+), \\
&\text{s.t. } V+\eta = pV_P^+ + (1-p)\hat{V}_+, \\
&\quad p_+ \in [0,1], \quad \hat{V}_+ \in \left[V_L^+, V_F^+\right].
\end{aligned} \tag{A.6}
$$

Let $\hat{V}$ be defined as $\hat{V}_+ - \eta$. Since $\hat{V}_+$ belongs to the interval $[V_L^+, V_F^+]$, $\hat{V}$ belongs to the interval

$$
\begin{aligned}
\left[V_L^+ - \eta, V_F^+ - \eta\right] \quad &= \left[(1-\beta)V_P^+ + \beta V_F^+ - \eta, V_F^+ - \eta\right] \\
&= \left[(1-\beta)\left(V_P^+ - \eta\right) + \beta\left(V_F^+ - \eta\right), V_F^+ - \eta\right] \\
&= \left[(1-\beta)V_P + \beta V_F, V_F\right] = [V_L, V_F].
\end{aligned} \tag{A.7}
$$

The third line makes use of the fact that

$$
V_P^+ - \eta = -\eta = V_P. \tag{A.8}
$$

The third line also makes use of the fact that

$$\begin{aligned} V_F^+ - \eta &= \int_{-\phi} (x+\phi) dF(x) - \eta \\ &= \int_{-\phi} x dF(x) + \phi(1 - F(-\phi)) - \eta \\ &= \int_{\eta} (x - \eta) dF(x) = V_F. \end{aligned} \tag{A.9}$$

Using the definition of $\hat{V}$ and the above observations, we can rewrite (A.6) as

$$\begin{aligned} J_+(V+\eta) &= \max_{p,\hat{V}} pJ_P^+ + (1-p)\hat{J}_+(\hat{V}+\eta) \\ &\text{s.t. } V+\eta = pV_P^+ + (1-p)(\hat{V}+\eta), \\ &\quad p \in [0,1], \quad \hat{V} \in [V_L, V_F],. \end{aligned} \tag{A.10}$$

Rearranging terms in the objective function and using the fact that $J_P^+ = J_P$, we can rewrite (A.10) as

$$\begin{aligned} J_\phi(V+\eta) &= \max_{p,\hat{V}} pJ_P + (1-p)\hat{J}_\phi(\hat{V}+\eta) \\ &\text{s.t. } V = pV_P + (1-p)\hat{V}, \\ &\quad p \in [0,1], \quad \hat{V} \in [V_\ell, V_F], . \end{aligned} \tag{A.11}$$

The value functions $\hat{J}_+(\hat{V}+\eta) = \hat{J}(\hat{V})$ and $J_+(V+\eta) = J(V)$ solve the functional equations (A.5) and (A.11), since $\hat{J}(\hat{V})$ and $J(V)$ are the solution to (2.4) and (2.1), (A.5) is identical to (2.4), and (A.11) is identical to (2.1). Parts 2, 3 and 4 of the proposition directly follow from the change in choice variables implemented in the derivation of (A.5) and (A.11). ■